

# Multiple-trait model by Bayesian inference applied to environment efficient *Coffea arabica* with low-nitrogen nutrient

Antônio Carlos da Silva Júnior<sup>1,\*</sup> , Waldênia de Melo Moura<sup>2</sup> , Lívia Gomes Torres<sup>1</sup> , Iara Gonçalves dos Santos<sup>1</sup> , Michele Jorge da Silva<sup>1</sup> , Camila Ferreira Azevedo<sup>3</sup> , Cosme Damião Cruz<sup>1</sup> 

1. Universidade Federal de Viçosa  – Departamento de Biologia Geral – Viçosa (MG), Brazil.

2. Empresa de Pesquisa Agropecuária de Minas Gerais  – Unidade Regional da Zona da Mata – Viçosa (MG), Brazil.

3. Universidade Federal de Viçosa  – Departamento de Estatística – Viçosa (MG), Brazil.

**Received:** Aug. 11, 2022 | **Accepted:** Jan. 19, 2023

**Section Editor:** Christian Cilas 

\***Corresponding author:** antonio.silva.c.junior@gmail.com

**How to cite:** Silva Júnior, A. C., Moura, W. M., Torres, L. G., Santos, I. G., Silva, M. J., Azevedo, C. F. and Cruz, C. D. (2023). Multiple-trait model by Bayesian inference applied to environment efficient *Coffea arabica* with low-nitrogen nutrient. *Bragantia*, 82, e20220157. <https://doi.org/10.1590/1678-4499.20220157>

**ABSTRACT:** Identifying *Coffea arabica* cultivars that are more efficient in the use of nitrogen is an important strategy and a necessity in the context of environmental and economic impacts attributed to excessive nitrogen fertilization. Although *Coffea arabica* breeding data have a multi-trait structure, they are often analyzed under a single trait structure. Thus, the objectives of this study were to use a Bayesian multitrait model, to estimate heritability in the broad sense, and to select arabica coffee cultivars with better genetic potential (desirable agronomic traits) in nitrogen-restricted cultivation. The experiment was carried out in a greenhouse with 20 arabica coffee cultivars grown in a nutrient solution with low-nitrogen content (1.5 mM). The experimental design used was in randomized blocks with three replications. Six agromorphological traits of the arabica coffee breeding program and five nutritional efficiency indices were used. The Markov Chain Monte Carlo algorithm was used to estimate genetic parameters and genetic values. The agromorphological traits were considered highly heritable, with a credibility interval (95% probability):  $H^2 = 0.9538 - 5.89E-01$ . The Bayesian multitrait model presents an adequate strategy for the genetic improvement of arabica coffee grown in low-nitrogen concentrations. Coffee arabica cultivars Icatu Precoce 3282, Icatu Vermelho IAC 4045, Acaiaí Cerrado MG 1474, Tupi IAC 1669-33, Catucaí 785/15, Caturra Vermelho and Obatã IAC 1669/20 demonstrated greater potential for cultivation in low-nitrogen concentration.

**Key words:** high performance, heritable, credibility interval.

## INTRODUCTION

Brazil is the world's largest producer and exporter of arabica coffee (*Coffea arabica* L.), responsible for the production of 47.7 million bags of processed coffee (Conab 2022). Coffee growing regions, in general, have naturally acidic soils with low fertility. Therefore, soil pH correction and the use of large amounts of chemical fertilizers are necessary to ensure the maximum productive potential of coffee trees. Among the important macronutrients for plant growth and development, nitrogen is the most required by coffee plants, as it performs fundamental biochemical functions, such as amino acid and protein synthesis, impacting photosynthesis, formation of flower buds, in addition to the effects on the chemical composition of the fruit (Clemente et al. 2015). Brazil is the fourth largest consumer of nitrogen fertilizers in the world, which makes it dependent on the import of these inputs (GloboFert 2022). Faced with the current fertilizer supply crisis, research focusing on mineral nutrition and the identification of more efficient genetic materials in the use of nutrients have been identified as an alternative to reducing the use of these inputs.

The selection of cultivars adapted to soils with low fertility is a challenge for genetic improvement programs. The evaluation of several morphological and physiological characters is necessary, which makes the evaluation and selection process difficult since superior cultivars combine ideal attributes for several characteristics simultaneously. Thus, statistical methodologies must be used to evaluate information with a multitrait structure, which correctly represents the genetic and phenotypic variation in the data (Malosetti et al. 2008).

Bayesian multitrait models are suitable for genetic evaluations in plants (Junqueira et al. 2016, Volpato et al. 2019, Silva Junior et al. 2022). In addition, these models allow the estimation of variance components and genetic values for each trait (Peixoto et al. 2021, Silva Junior et al. 2022), jointly modeling multiple traits. Mora and Serra (2014), Junqueira et al. (2016), Torres et al. (2018), Volpato et al. (2019) and Silva Junior et al. (2022) demonstrated the potential of the Bayesian approach for genetic evaluation in plant breeding, considering multi-environment and multi-trait models. However, there is still a lack of information from multi-trait models under a Bayesian approach for the cultivation of arabica coffee in environments with low fertility.

The objectives of this study were to estimate genetic parameters of arabica coffee grown under low-nitrogen conditions using a Bayesian multi-trait model and to select arabica coffee cultivars with better genetic potential.

## MATERIALS AND METHODS

### Field experiments

The experiment was carried out in the state of Minas Gerais, Brazil, by the Agricultural Research Company of Minas Gerais/Southeast (EPAMIG/Southeast), in a greenhouse located at the experimental field Diogo Alves de Melo, in the Universidade Federal de Viçosa (20° 45' S, 42° 52' W, 648 m).

Twenty *Coffea arabica* cultivars (Table 1) were evaluated for different agromorphological traits in an aerated static nutrient solution containing low nitrogen (1 mmol·L<sup>-1</sup>).

**Table 1.** List of cultivars and genealogy of *Coffea arabica* L. evaluated in under low nitrogen in a greenhouse condition.

Cultivars	Genealogy
1 Catuaí Amarelo IAC 62	Caturra Amarelo IAC 476-11 × Mundo Novo IAC 374-19
2 Paraíso MG H 419-1	Caturra Amarelo IAC 30 × Híbrido de Timor UFV 445-46
3 Icatu Vermelho IAC 4045	<i>Coffea canephora</i> × Bourbon Vermelho
4 Obatã IAC 1669/20	Villa Sarchi × Híbrido de Timor
5 Caturra Amarelo	Mutação do Caturra Vermelho
6 IPR 102	Icatu × (Catuaí Amarelo IAC 66 × Catuaí Vermelho IAC 99)
7 Catuaí Vermelho IAC 15	Caturra Amarelo IAC 476-11 × Mundo Novo
8 Rubi MG 1192	Catuaí × Mundo Novo
9 IPR 103	Icatu × (Catuaí Amarelo IAC 66 × Catuaí Vermelho IAC 99)
10 Araponga MG1	Caturra Amarelo IAC 86 × Híbrido de Timor UFV 446-08
11 Topázio MG 1190	Catuaí Amarelo × Novo Mundo
12 San Ramon	Progênie C 1034-4 Introdução da Costa Rica
13 Icatu Precoce 3282	Icatu Vermelho × Bourbon Amarelo
14 Tupi IAC 1669-33	Villa Sarchi × Híbrido de Timor
15 Catucaí 785/15	Icatu Vermelho × Catuaí Vermelho
16 Acaiá Cerrado MG 1474	Seleção de Mundo Novo
17 Oeiras MG 6851	Híbrido CIFC HW 26/5 (Caturra Vermelho × Híbrido de Timor)
18 São Bernardo	Mutação Típica do Café
19 Pau Brasil MG1	Catuaí Vermelho IAC 15 × Híbrido de Timor UFV 442/34
20 Caturra Vermelho	Mutação de Bourbon Vermelho

The experiment was carried out in a randomized block design with three replicates. The plots consisted of two plants grown in pots with nutrient solution. The cultivars were sown in a sand bed sterilized with hydrochloric acid (HCl) ( $0.1 \text{ mol}\cdot\text{dm}^{-3}$ ). At 120 days, two seedlings at the cotyledonary leaf stage were transplanted into pots containing 8 L of nutrient solution (Hoagland and Arnon 1950). The nutrient solution was completed weekly with deionized water, and the pH was adjusted between 5.5 and 6.5 using HCl ( $0.1 \text{ mol}\cdot\text{dm}^{-3}$ ) and sodium hydroxide (NaOH) ( $0.1 \text{ mol}\cdot\text{dm}^{-3}$ ). The electrical conductivity (EC) was monitored by the change of the nutrient solution when its depletion reached 30% of the initial EC.

At 168 days after transplanting, the following morpho-agronomic traits were evaluated: stem diameter (SD, mm), measured with a caliper at 5 cm from the stem base; plant height (PH, cm), measured from the base of the orthotropic branch to the plant apex; internode length (IL, cm), calculated as a ration between plant height and node number; number of leaf pairs (NLP), obtained by counting the whole plant; number of nodes (NNO), obtained by counting the nodes in the main branch (orthotropic); leaf area (LA,  $\text{dm}^2$ ), quantified after harvesting using the leaf area meter model AT Delta-TDevices.

The plants were sectioned into roots, stem, and leaves, dried in an oven with forced air circulation at  $70 \text{ }^\circ\text{C}$  for 72 h and weighted to get: root dry matter (MSR), stem dry matter (MSC), and leaf dry matter (MSF). Shoot mass (MSPA) consisted of the sum of MSF and MSC, while the total dry mass (MST) was obtained by the sum of MSPA and MSR. The dried plant material was ground in a “Willey” mill to determine the nitrogen content according to the protocol of Empresa Brasileira de Pesquisa Agropecuária (Embrapa 2009). Nitrogen content was estimated as the product between nutrient content in different parts of the plants and the total dry mass.

Nutritional efficiency indexes were estimated as described by Siddiqi and Glass (1981) and Bailian et al. (1991): rooting efficiency ( $\text{EE}, \text{g}^2/\text{mg}$ ) = (root dry matter)<sup>2</sup>/total N content; absorption efficiency ( $\text{EA}, \text{mg}/\text{g}$ ) = total N content/root dry matter; translocation efficiency ( $\text{ET}, \text{mg}/\text{g}$ ) = shoot N content/total N content; utilization efficiency ( $\text{EU}, \text{g}^2/\text{mg}$ ) = (MST)<sup>2</sup>/total N content; biomass production efficiency ( $\text{EPB}, \text{g}^2/\text{mg}$ ) = (MSPA)<sup>2</sup>/shoot N content.

## Biometric analysis

The data was analyzed using the univariate and multi-trait models through Markov Chain Monte Carlo (MCMC) Bayesian approach. The multi-trait model was given by Eq. 1:

$$y = X\beta + Zg + \varepsilon \quad (1)$$

in which:  $y$  = the vector of phenotypic data.

The conditional distribution was given by Eq. 2:

$$y | \beta, g, i, G, R \sim N(X\beta + Zg, R \otimes I) \quad (2)$$

$G$  = the matrix of genotypic covariance;  $R$  = the matrix of residual covariance;  $I$  = an identity matrix;  $\beta$  = vector of systematic effects (genotypes mean and replication effects), assumed as  $\beta \sim N(\beta, \Sigma\beta \otimes I)$ ;  $g$  = the vector of genotype effects, assumed as  $g | G, \sim N(0, G \otimes I)$ ;  $e$  = the vector of residuals, assumed as  $e | R, \sim N(0, R \otimes I)$ .

The uppercase bold letters  $X$  and  $Z$  refer to the incidence matrices for the effects  $\beta$  and  $g$ , respectively. The R package *MCMCglmm* (Hadfield 2010) was used to fit the model.

We assumed that  $G$  and  $R$  follow an inverted Wishart distribution  $WI(v, V)$ , with hyperparameters  $v$  and  $V$  (Sorensen and Gianola 2002). Hyperparameters for all prior distributions have been selected to provide non-informative or flat prior distributions. For the systematic effect ( $\beta$ ), a uniform distribution was assigned. In addition, the parameters  $\beta$ ,  $g$ ,  $G$ , and  $R$  were estimated following the set posterior distribution:  $P(\beta, g, G, R | y) \propto P(y | \beta, g, G, R) \times P(\beta, g, G, R)$ .

In total, 1,800,000 samples were generated. A burn-in of 10,000 and thin of 10 iterations were assumed, resulting in a total of 1,790,000 samples. The convergence of the MCMC was verified by the criterion of Geweke (1992), using the R packages *boa* (Smith 2007) and convergence diagnosis and output analysis (CODA) (Plummer et al. 2006).

The model was compared using the deviation information criterion (DIC) proposed by Spiegelhalter et al. (2002) (Eq. 3):

$$DIC = D(\bar{\theta}) + 2p_D \quad (3)$$

in which:  $D(\bar{\theta})$  = a point estimate of the deviance obtained by replacing the parameters with their posterior means estimates in the likelihood function;  $p_D$  = the effective number of model parameters.

Models with a lower DIC should be preferred over models with a higher DIC.

Variance components, broad-sense heritability ( $H^2$ ), genotypic correlation coefficients between traits, and breeding values were calculated from the posterior distribution. Intervals of higher posterior density (HPD) were estimated for all traits using the R package *boa* (Smith 2007). *A posteriori* estimates of  $H^2$  for each trait and each iteration were calculated from the later samples of variance components, using Eq. 4:

$$H^{2(i)} = \frac{\sigma_g^{2(i)}}{(\sigma_g^{2(i)} + \sigma_r^{2(i)})} \quad (4)$$

in which:  $\sigma_g^{2(i)}$  = the genetic variance component of each iteration;  $\sigma_r^{2(i)}$  = the residual variance component of each iteration.

### Selection based on selection index

The multi-trait index based on factor analysis and genotype-ideotype distance (FAI-BLUP) (Rocha et al. 2018) was used to identify superior coffee genotypes under low-nitrogen. The formula is as Eq. 5:

$$P_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{i=1; j=1}^{i=n; j=m} \frac{1}{d_{ij}}} \quad (5)$$

in which:  $P_{ij}$  = probability of the  $i^{\text{th}}$  genotype ( $i = 1, 2, \dots, n$ ) to be similar to the  $j^{\text{th}}$  ideotype ( $j = 1, 2, \dots, m$ );  $d_{ij}$  = genotype-ideotype distance from  $i^{\text{th}}$  genotype to  $j^{\text{th}}$  ideotype, based on standardized mean distance.

Selection gains (SG) were estimated from the FIA-BLUP considering three different selection intensities: 35, 50, and 60%, which refers to the selection of seven, ten, and 12 genotypes, respectively, as follows (Eq. 6):

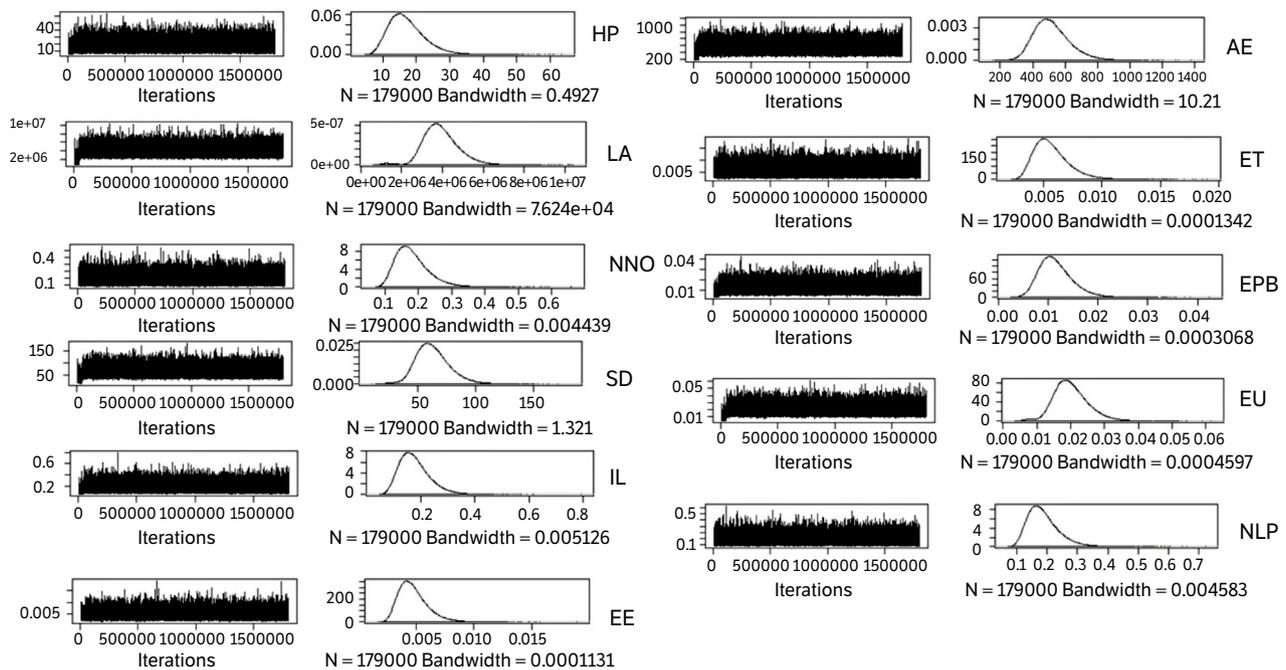
$$SG(\%) = \left( \frac{X_s - X_0}{X_0} \right) \quad (6)$$

in which:  $X_s$  = the mean of the selected genotypes;  $X_0$  = the overall population mean.

## RESULTS AND DISCUSSION

Geweke's criterion indicated convergence for all dispersion parameters generating 1,800,000 MCMC iterations, 10,000 samples for burn-in, and a sampling interval of 10, totaling 1,790,000 effective samples used to estimate the variance components (Fig. 1). All chains [components of (co)variance] reached convergence by this criterion. Similar posterior averages were obtained for the variance components, suggesting normal-appearing density. The DIC suggests that the full model for multi-trait is the one that best fits the data, which reveals the significance of genotypic effects (DIC = 1123.39 for the full model and 1,414.31 for the restricted one). This is justified by the lowest DIC value of the full model. Spiegelhalter et al. (2002) suggest that the use of the complete model can lead to a greater prediction in the estimation of parameters.

Subsequent mean estimates for the variance components suggested  $\chi^2$  density and normal distributions (Fig. 1). Thus, it is possible to observe that all the characteristics presented a  $\chi^2$  distribution (of which the Wishart distribution is a generalization).



PH: plant height; SD: stem diameter; NNO: number of nodes; IL: internode length; NLP: number of leaf pairs; LA: leaf area; EE: rooting efficiency; EA: absorption efficiency; ET: translocation efficiency; EU: efficiency of use; EPB: biomass production efficiency.

**Figure 1.** Convergence for the genotypic variance of the 11 traits analyzed in the multi-trait model. The posterior density and genetic variance estimates are on both the right and the left.

The  $H^2$  estimates were different for the mean and posterior density range (HPD) (Table 2). The highest  $H^2$  values were observed for the PH and IL traits (greater than 80%). On the other hand, the lowest estimates were for the EA and LA traits. The low heritability observed estimate depends on the number of genotypes evaluated. The Bayesian approach used can be recommended for situations involving small sample sizes space (Torres et al. 2018, Silva Junior et al. 2022). The PH, NLP, NNO, IL, and EE traits were considered highly heritable, with a credibility interval (95% probability) ranging from 0.6800 to 0.9538; 0.5890 to 0.9326; 0.5920 to 0.9313; 0.6490 to 0.9494; and 0.6010 to 0.9224, respectively (Table 2).  $H^2$  estimates greater than 70% for these same traits were also found in coffee in a potassium-restricted crop, using analysis of variance (ANOVA) (Moura et al. 2016).

**Table 2.** Later inferences for mean and posterior density range (HPD 95%) of broad-sense heritability ( $H^2$ ) considering the multi-trait model.

Trait	$H^2$	HPD 95%	
		Lower Bound	Upper Bound
PH	0.83	0.6800	0.9538
NLP	0.77	0.5890	0.9326
NNO	0.77	0.5920	0.9313
SD	0.14	0.00350	0.3286
LA	0.02	6.23E-09	0.0023
IL	0.81	0.6490	0.9494
EE	0.77	0.6010	0.9224
EA	0.04	7.47E-06	0.1626
ET	0.72	0.5290	0.8982
EPB	0.48	0.2400	0.7396
EU	0.49	0.0366	0.12525

PH: plant height; SD: stem diameter; NNO: number of nodes; IL: internode length; NLP: number of leaf pairs; LA: leaf area; EE: rooting efficiency; EA: absorption efficiency; ET: translocation efficiency; EU: efficiency of use; EPB: biomass production efficiency.

The multi-trait Bayesian model has been successfully used in several crops, such as flood-irrigated rice, where  $H^2$  estimates were higher than 80% (Silva Junior et al. 2022), and in maize lines, where the heritability for nitrogen use efficiency was 50%, considered highly heritable (Torres et al. 2018). These authors reported that the Bayesian multi-trait model makes estimates more accurate than in individual models due to taking into account the correlation between the traits. Mora et al. (2019) evaluated *E. globulus* clones and found moderate to high heritability values for tree heights ranging from 12 to 41% (mode value of the posterior distribution of heritability).

In addition to the statistical model, the heritability of a trait is crucial to improving the prediction (Lorenz et al. 2011, Gill et al. 2021). Low heritability estimates result in lower accuracy in predicting individual trait (Heffner et al. 2009). The application of multi-trait models, in turn, can improve the prediction of poorly heritable characters using information from correlated characters that have high heritability (Jia and Jannink 2012, Jiang et al. 2015, Lado et al. 2018, Gill et al. 2021; Bhatta et al. 2020).

Jia and Jannink (2012) also indicated that a multi-trait model is more effective when the genetic correlation between traits is moderate. Guo et al. (2020) reported that characters with lower heritability performed better than those with high heritability through the multi-trait model, as it contemplates the interaction between traits  $\times$  genotypes, and provides a better estimate of the correlation between characters. Schulthess et al. (2018) and Montesinos-López et al. (2018) showed that the performance of multi-trait analysis depends considerably on the number of missing characters in only some individuals or all individuals. Precise estimates of genetic parameters bring new perspectives on the application of Bayesian methods to solve modeling problems in the genetic improvement of arabica coffee for cultivation under low-nitrogen concentration. This is justified by the parameters that lie within the posterior density range (HPD 95%).

The estimates of genetic variance and repetition for each iteration were discrepant for the mean, median, and mode in each character (Table 3). The lowest values were observed for traits of efficiency indexes, NNO and NLP. On the other hand, PH and LA presented the highest estimates. Similar results were reported by Moura et al. (2016) for arabica coffee under potassium limiting conditions, using ANOVA. The multi-trait model used in the present work showed great performance in estimating genetic and residual variances since the estimates are within the posterior density range (HPD 95%). This model presents credibility intervals that are more accurate when compared to the confidence intervals obtained in frequentist inference (Gazola et al. 2016).

**Table 3.** Genetic and residual variance of 11 traits of arabica coffee cultivars in low-nitrogen cultivation, using multi-trait models.

Trait	$\sigma_g^2$					$\sigma_r^2$				
	Mean	Median	Mode	Lower Bound	Upper Bound	Mean	Median	Mode	Lower Bound	Upper Bound
PH	95.02	82.42	65.50	28.14	191.51	17.21	16.37	15.03	28.15	191.51
NLP	0.75	0.64	0.65	0.19	1.57	0.17	0.18	0.16	0.19	1.57
NNO	0.74	0.63	0.49	0.19	1.53	0.17	0.18	0.16	0.19	1.53
SD	10.80	7.87	5.52	0.10	27.73	63.40	61.76	58.03	0.10	27.73
LA	87,463.57	108.58	79.55	0.0202	888,000	3,913,594	3,834,606	3,685,159	0.0202	888,000
IL	0.46	0.76	0.58	0.26	1.76	0.18	0.17	0.15	0.26	1.76
EE	0.018	0.016	0.013	0.006	0.004	0.005	0.004	0.004	0.006	0.004
EA	27.22	9.06	1.13	0.0046	93.96	520.86	507.48	481.43	0.0046	93.96
ET	0.016	0.014	0.012	0.005	0.030	0.006	0.006	0.005	0.005	0.030
EPB	0.011	0.010	0.007	0.003	0.020	0.012	0.010	0.010	0.003	0.020
EU	0.006	0.007	0.004	0.002	0.014	0.020	0.019	0.018	0.002	0.014

PH: plant height; SD: stem diameter; NNO: number of nodes; IL: internode length; NLP: number of leaf pairs; LA: leaf area; EE: rooting efficiency; EA: absorption efficiency; ET: translocation efficiency; EU: efficiency of use; EPB: biomass production efficiency;  $\sigma_g^2$ : the variance genetic;  $\sigma_r^2$ : the replication for each iteration.

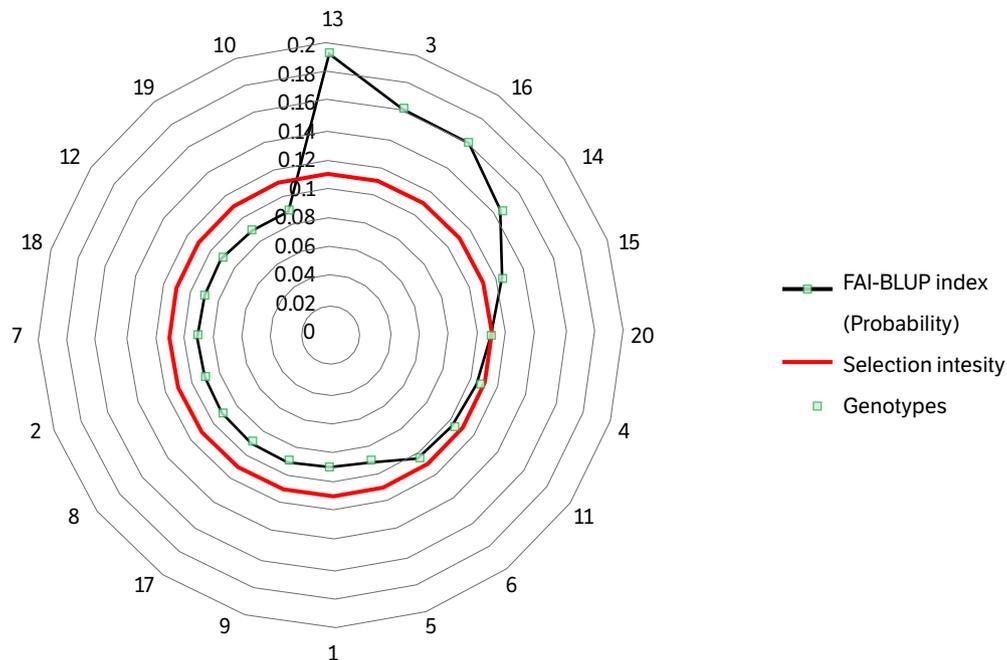
The selection gains obtained by the FAI-BLUP index considering three different selection intensities: 35, 50, and 60%, which refers to the selection of seven, 10, and 12 genotypes, for 11 traits of arabica coffee cultivars in an efficient low-nitrogen environment, is represented in Table 4. The FAI-BLUP index indicated discrepant selection gains between different selection intensities for the same character (Table 4). Selection gains increased with increasing selection intensity. The highest GS was estimated for PH and IL regardless of selection intensity. On the other hand, the N absorption efficiency showed gains for all selection intensities.

**Table 4.** Percentage of selection gains, factor number, and commonalities obtained by the FIA-BLUP index considering three different selection intensities: 35, 50, and 60%, which refers to the selection of seven, 10, and 12 genotypes, for eleven traits of arabica coffee cultivars in an efficient low-nitrogen environment.

Trait	Factor	Commonalities	7	10	12
PH	1	0.96	14.64	7.39	4.89
SD	1	0.96	2.75	1.98	1.54
IL	1	0.95	13.45	7.57	5.38
EA	1	0.98	0.00	0.00	0.00
NLP	2	0.68	0.90	-0.45	-0.73
NNO	2	0.98	0.67	-0.49	-0.71
LA	2	0.71	-0.02	0.06	0.09
EE	2	0.98	-0.04	-0.28	-1.24
ET	2	0.85	1.86	1.03	0.72
EPB	3	0.98	8.27	5.17	3.64
EU	3	0.97	7.25	4.40	2.88

PH: plant height; SD: stem diameter; NNO: number of nodes; IL: internode length; NLP: number of leaf pairs; LA: leaf area; EE: rooting efficiency; EA: absorption efficiency; ET: translocation efficiency; EU: efficiency of use; EPB: biomass production efficiency.

Figure 2 represents the classification of the 20 arabica coffee cultivars, considering the evaluated traits, according to the FAI-BLUP index and its associated spatial probability. Under a selection intensity of 35%, the cultivars with the highest genetic potential at low-nitrogen concentration were Icatu Precoce 3282 (13), Icatu Vermelho IAC 4045 (3), Acaiá Cerrado MG 1474 (16), Tupi IAC 1669-33 (14), Catucaí 785/15 (15), Caturra Vermelho (20), and Obatã IAC 1669/20 (4). Rocha et al. (2018) used the FAI-BLUP index to select elephant grass genotypes and claim that this index classifies genotypes based on multi-trait free of multicollinearity.



**Figure 2.** Selection considering 35% of selection intensity (selection of seven cultivars). The green dotted line indicates the arabica coffee genotypes evaluated in a low-nitrogen environment. The cultivars selected by the FAI-BLUP index correspond to the green points outside the red line.

## CONCLUSION

The Bayesian multi-trait model is an adequate strategy for the genetic improvement of arabica coffee grown under low-nitrogen concentration.

Arabica coffee cultivars Icatu Precoce 3282, Icatu Vermelho IAC 4045, Acaia Cerrado MG 1474, Tupi IAC 1669-33, Catucaí 785/15, Caturra Vermelho, and Obatã IAC 1669/20 have greater genetic potential for cultivation in low-nitrogen concentration.

## AUTHORS' CONTRIBUTION

**Conceptualization:** Silva Júnior, A. C., Moura, W. M., Cruz, C. D. and Azevedo, C. F.; **Methodology:** Silva Júnior, A. C., Torres, L. G., Cruz, C. D. and Azevedo, C. F.; **Investigation:** Silva Júnior, A. C., Moura, W. M., Cruz, C. D. and Azevedo, C. F.; **Writing – Original Draft:** Silva Júnior, A. C., Moura, W. M., Santos, I. G., Silva, M. J., Torres, L. G., Cruz, C. D. and Azevedo, C. F.; **Writing – Review and Editing:** Silva Júnior, A. C., Moura, W. M., Santos, I. G., Silva, M. J., Torres, L. G., Cruz, C. D. and Azevedo C. F.; **Funding Acquisition:** Silva Júnior, A. C., Moura, W. M., Cruz, C. D. and Azevedo, C. F.; **Supervision:** Silva Júnior, A. C., Moura, W. M., Cruz, C. D. and Azevedo, C. F.

## DATA AVAILABILITY STATEMENT

The data belong to the Agricultural Research Company of Minas Gerais, in Brazil. There are ethical restrictions on sharing the used data set because this one still contains important data about the genotype information; in addition, the data os owned by state-owned organization. However, data access requests may be directed to: epamigsudeste@epamig.br.

## FUNDING

Fundação de Amparo à Pesquisa do Estado de Minas Gerais

<https://doi.org/10.13039/501100004901>

Conselho Nacional de Desenvolvimento Científico e Tecnológico

<https://doi.org/10.13039/501100003593>

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

<https://doi.org/10.13039/501100002322>

Financial Code 001

Coffee Research Consortium (CP-Café)

## ACKNOWLEDGMENTS

The authors would like to thank Prof. Dr. Fabyano Fonseca e Silva (*in memoriam*) and Dr. Paulo César de Lima (*in memoriam*).

## REFERENCES

- Bailian, L., Mckeand, S. E. and Allen, H. L. (1991). Genetic variation in nitrogen use efficiency of lobeolly pine seedlings. *Forest Science*, 37, 613-626. <https://doi.org/10.1093/forestscience/37.2.613>
- Bhatta, M., Gutierrez, L., Cammarota, L., Cardozo, F., Germán, S., Gómez-Guerrero, B., Pardo, M. F., Lanaro, V., Sayas, M. and Castro, A. J. (2020). Multi-trait genomic rediction model increased the predictive ability for agronomic and malting quality traits in barley (*Hordeum vulgare* L.). *G3 Genes Genomes Genetics*, 10, 1113-1124. <https://doi.org/10.1534/g3.119.400968>
- Clemente, J. M., Martinez, H. E. P., Alves, L. C., Finger, F. L. and Cecon, P. R. (2015). Effects of nitrogen and potassium on the chemical composition of coffee beans and on beverage quality. *Acta Scientiarum. Agronomy*, 37, 297-305. <https://doi.org/10.4025/actasciagron.v37i3.19063>
- [CONAB] Companhia Nacional de Abastecimento (2022). Acompanhamento da safra brasileira de café. Brasília: Companhia Nacional de Abastecimento.
- [EMBRAPA] Empresa Brasileira de Pesquisa Agropecuária (2009). Manual de análises químicas de solos, plantas e fertilizantes. 2st ed. Brasília: Embrapa Informação Tecnológica.
- Gazola, S., Scapim, C. A., Araujo, Â. M. M., Rossi, R. M., Amaral, A. T. and Vivas, M. (2016). Nonlinear models to describe the maize seed quality during the maturation stage: a Bayesian approach. *Australian Journal of Crop Science*, 10, 598-603. <https://doi.org/10.21475/ajcs.2016.10.05.p6361>
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (Eds.), *Bayesian Statistics 4* (p. 625-631). Oxford: University Press.
- Gill, H. S., Halder, J., Zhang, J., Brar, N. K., Rai, T. S., Hall, C., Bernardo, A., Amand, P. S., Bai, G., Olson, E., Ali, S., Turnipseed, B. and Sehgal, S. K. (2021). Multi-Trait multi-environment genomic prediction of agronomic traits in advanced breeding lines of winter wheat. *Frontiers in Plant Science*, 12, 709545. <https://doi.org/10.3389/fpls.2021.709545>
- GloboFert. Available at: <https://globalfert.com.br/>. Accessed on: Mar. 20, 2022.
- Guo, J., Khan, J., Pradhan, S., Shahi, D., Khan, N., Avci, M., Mcbreen, J., Harrison, S., Brown-Guedira, G., Murphy, J. P., Johnson, J., Mergoum, M., Mason, R. E., Ibrahim, A. M. H., Sutton, R., Griffey, C. and Babar, M. A. (2020). Multi-trait genomic prediction of yield-related traits in US soft wheat under variable water regimes. *Genes*, 11, 1270. <https://doi.org/10.3390/genes11111270>
- Hadfield, J. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R Package. *Journal Stat Software*, 33, 1-22. <https://doi.org/10.18637/jss.v033.i02>
- Heffner, E. L., Sorrells, M. E. and Jannink J. L. (2009). Genomic selection for crop improvement. *Crop Science*, 49, 1-12. <https://doi.org/10.2135/cropsci2008.08.0512>
- Hoagland, D. R. and Arnon, D. I. (1950). The water culture method for growing plants without soils. Berkeley: California Agricultural Experimental Station.
- Jia, Y. and Jannink, J. L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics*, 192, 1513-1522. <https://doi.org/10.1534/genetics.112.144246>
- Jiang, J., Zhang, Q., Ma, L., Li, J., Wang, Z. and Liu, J. F. (2015). Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. *Heredity* 115, 29–36. <https://doi.org/10.1038/hdy.2015.9>
- Junqueira, V. S., Azevedo, P. L., Laviola, B. G., Bhering, L. L., Mendonça, S., Costa, T. S. A. and Antoniassi, R. (2016). Correction: Bayesian multi-trait analysis reveals a useful tool to increase oil concentration and to decrease toxicity in *Jatropha curcas* L. *PLoS One*, 11, e0161046. <https://doi.org/10.1371/journal.pone.0161046>

- Lado, B., Vázquez, D., Quincke, M., Silva, P., Aguilar, I. and Gutiérrez, L. (2018). Resource allocation optimization with multi-trait genomic prediction for bread wheat (*Triticum aestivum* L.) baking quality. *Theoretical and Applied Genetics*, 131, 2719-2731. <https://doi.org/10.1007/s00122-018-3186-3>
- Lorenz, A. J., Chao, S., Asoro, F. G., Heffner, E. L., Hayashi, T., Iwata, H., Smith, K. P., Sorrells, M. E. and Jannink, J. L. (2011). Genomic selection in plant breeding: knowledge and prospects. *Advances in Agronomy*, 110, 77-123. <https://doi.org/10.1016/B978-0-12-385531-2.00002-5>
- Malosetti, M., Ribaut, J. M., Vargas, M., José Crossa, J. and van Eeuwijk, F. A. (2008). A multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea mays* L.). *Euphytica*, 161, 241-257. <https://doi.org/10.1007/s10681-007-9594-0>
- Montesinos-López, O. A., Montesinos-López, A., Montesinos-López, J. C., Crossa, J., Luna-Vázquez, F. J. and Salinas-Ruiz, J. (2018). A Bayesian multiple-trait and multiple-environment model using the matrix normal distribution. *Physical Methods for Stimulation of Plant and Mushroom*, 19. <https://doi.org/10.5772/intechopen.71521>
- Mora, F. and Serra, N. (2014). Bayesian estimation of genetic parameters for growth, stem straightness, and survival in *Eucalyptus globulus* on an Andean Foothill site. *Tree Genetics & Genomes*, 10, 711-719. <https://doi.org/10.1007/s11295-014-0716-2>
- Mora, F., Ballesta, P. and Serra, N. (2019). Bayesian analysis of growth, stem straightness and branching quality in full-sib families of *Eucalyptus globulus*. *Bragantia*, 78, 328-336. <https://doi.org/10.1590/1678-4499.20180317>
- Moura, W. M., Soares, Y. J. B., Amaral Júnior, A. T., Gravina, G. A., Barili, L. D. and Vieira, H. D. (2016). Biometric analysis of arabica coffee grown in low potassium nutrient solution under greenhouse conditions. *Genetic Molecular Research*, 15, gmr.15038753. <https://doi.org/10.4238/gmr.15038753>
- Peixoto, M. A., Evangelista, J. S. P. C., Coelho, I. F., Alves, R. S., Laviola, B. G., Fonseca e Silva F., Resende, M. D. V. and Bhering, L. L. (2021). Multiple-trait model through Bayesian inference applied to *Jatropha curcas* breeding for bioenergy. *PLoS One*, 16, e0247775. <https://doi.org/10.1371/journal.pone.0247775>
- Plummer, M., Best, N., Cowles, K., Vines, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6, 7-11.
- Rocha, J. R. A. S. C., Machado, J. C. and Carneiro, P. C. S. (2018). Multitrait index based on factor analysis and ideotype design: proposal and application on elephant grass breeding for bioenergy. *GCB Bioenergy*, 10, 52-60. <https://doi.org/10.1111/gcbb.12443>
- Schulthess, A. W., Zhao, Y., Longin, C. F. H. and Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 131, 685-701. <https://doi.org/10.1007/s00122-017-3029-7>
- Siddiqi, M. Y. and Glass, A. D. M. (1981). Utilization index: a modified approach to the estimation and comparison of nutrient utilization efficiency in plants. *Journal of Plant Nutrition*, 4, 289-302. <https://doi.org/10.1080/01904168109362919>
- Silva Junior, A. C., Sant'Anna, I. C., Silva Siqueira, M. J., Cruz, C. D., Azevedo, C. F., Nascimento, M. and Soares, P. C. (2022). Multi-trait and multi-environment Bayesian analysis to predict the G x E interaction in flood-irrigated rice. *PLoS One*, 17, e0259607. <https://doi.org/10.1371/journal.pone.0259607>
- Smith, B. J. (2007). boa: an R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, 21, 1-37. <https://doi.org/10.18637/jss.v021.i11>
- Sorensen, D. A. and Gianola, D. (2002). Likelihood, Bayesian and MCMC methods in quantitative genetics: statistics for biology and health. New York: Springer-Verlag.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society. Statistical Methodology. Series B*, 64, 583-639. <https://doi.org/10.1111/1467-9868.00353>

Torres, L. G., Rodrigues, M. C., Lima, N. L., Trindade, T. F. H., Silva, F. F., Azevedo, C. F. and Lima, R. O. (2018). Multi-trait multi-environment Bayesian model reveals G x E interaction for nitrogen use efficiency components in tropical maize. *PLoS One*, 13, e0199492. <https://doi.org/10.1371/journal.pone.0199492>

Volpato, L., Alves, R. S., Teodoro, P. E., Resende, M. D. V., Nascimento, M., Nascimento A. C. C., Ludke, W. H., Silva, F. L. and Borém, A. (2019). Multi-trait multi-environment models in the genetic selection of segregating soybean progeny. *PLoS One*, 14, e0215315. <https://doi.org/10.1371/journal.pone.0215315>