



## SELECTION BETWEEN AND WITHIN SEGREGATING SOYBEAN POPULATIONS VIA SIMULATED INDIVIDUAL BLUP (BLUPIS)

André Ricardo Gomes Bezerra<sup>1\*</sup>, Felipe Lopes da Silva<sup>2</sup>, Tuneo Sedyama<sup>2</sup>,  
 Cosme Damião Cruz<sup>3</sup>, Amilton Ferreira da Silva<sup>4</sup>, Francisco Charles do  
Santos Silva<sup>5</sup>, Daniele Piano Rosa<sup>2</sup>

*1 Department of Plant Science, Fundação MS Para Pesquisa e Difusão de Tecnologias Agropecuárias, Maracaju, Mato Grosso do Sul, Brazil;*

*2 Department of Agronomy, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil.*

*3 Department of General Biology, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil;*

*4 Department of Agrarian Sciences, Universidade Federal de São João del-Rei, Sete Lagoas, Minas Gerais, Brazil;*

*5 Department of Agronomy, Universidade Estadual do Maranhão, Balsas, Maranhão, Brazil.*

*\* Corresponding author: André Ricardo Gomes Bezerra ([andrebezerra@fundacaoms.org.br](mailto:andrebezerra@fundacaoms.org.br)).*

**Abstract:** The use of methods with high statistical accuracy in the selection of genotypes can provide substantial gains to breeding programs. This study evaluated the suitability of the best linear unbiased prediction at individual level (BLUPI) with simulated individual BLUP (BLUPIS) procedure to select individuals in segregating soybean populations. Fifteen segregating populations were evaluated in the field during the 2014/15 crop season. Data for grain yield (g/plant) and growth cycle in days (divided in early and late) were collected and analyzed via restricted maximum likelihood/best linear unbiased prediction (REML/BLUP). In general, 67–92% of the individuals were selected in the three best populations. The efficiency of BLUPIS relative to BLUPI was 4.66 and 1.46 for grain yield and growth cycle, respectively. There was strong correlation between BLUPIS and BLUPI for grain yield. The selection by BLUPIS was more efficient than BLUPI for all the traits studied and contributes to the selection of individuals in representative numbers of the best populations, thereby providing substantial gain for soybean breeding programs.

**Keywords:** *Glycine max*, individual selection, REML/BLUP, grain yield.

## Introduction

In soybean breeding, genotype selection is based on phenotypic values, which confuse genotypic and environmental values. Thus, the use of methods with high statistical accuracy is an alternative, especially when working with traits with wide genetic variability (Resende and Alves 2020). The use of the restricted maximum likelihood/best linear unbiased prediction (REML/BLUP) is the standard procedure for the estimation of variance components, prediction of genetic parameters and optimal selection in several species (Resende, 2016). This approach allows accurate and unbiased prediction of genetic values, providing specific information on the merit of progeny and individuals in the progenies, with an advantage for unbalanced experiments. Furthermore, it even considers the environment of the repetition where the individual is located (Resende, 2007).

The procedure REML/BLUP has been widely used in the breeding of perennial crops, such as eucalyptus (Alves et al., 2018), coffee (Pereira et al., 2013) and *Jatropha curcas* (Spinelli et al., 2015). Despite having less frequent use in breeding annual species, the method has been used in crops, such as peanuts (Luz et al., 2010), triticale (Gowda et al., 2013), wheat (Pimentel et al., 2014) and popcorn (Viana et al., 2014).

Crossbreeding among higher parentals, followed by the selection of the best segregating populations, with or without the selection of individuals, is a classic procedure in soybean breeding. Due to the large number of populations evaluated annually, the breeder should focus on those populations markedly higher. The optimal selection strategy would be through genotypic values predicted by individual BLUP (BLUPI) that can use both family and individual level information for selection (Resende, 2002). However, information about the individuals is not usually obtained when soybean segregating populations are being evaluated, because these are estimated by total harvest of plots.

Under these conditions, a practical procedure of selection proposed by Resende and Barbosa (2006) for sugarcane could be useful in improving soybean to maximize the selection gain. The selection procedure called simulated individual BLUP (BLUPIS) aims to promote the

dynamic determination of the number of individuals selected per family, considering BLUP as a genotypic family effect. This strategy does not depend on the evaluation of individual plants within the plot, but rather indicates the number of individuals to be selected for each population, the total number of individuals to be evaluated and the number of families involved with the selected individuals. Given the above, this study evaluated the suitability of the BLUPIS compared to BLUPI to select individuals within segregating soybean populations.

## Material and Methods

### Obtaining phenotypic data

This study obtained 15 F<sub>2</sub> populations from biparental crosses that were performed by the Soybean Breeding Program of the Universidade Federal de Viçosa, Minas Gerais, Brazil, during the 2013/14 crop season. The experiment was conducted in the Experimental Field Diogo Alves de Melo, located at the University Campus, Viçosa, Minas Gerais, during the 2014/15 crop season. The region is geographically located at 20° 46' S latitude and 42° 52' W longitude. According to Köppen's classification, the climate of this region is type Cwa, i.e., humid temperate climate with a dry winter. The average rainfall is less than 60 mm in at least 1 month of this season. The summer is hot, with the average temperature of the hottest month over 20 °C.

The treatments were represented by 15 segregant populations obtained from complete balanced diallel crossbreeding and their six parental lines (Table 1). The soil was fertilized with 200 kg ha<sup>-1</sup> of a formula containing 4% N, 14% P<sub>2</sub>O<sub>5</sub> and 8% K<sub>2</sub>O.

The sowing was conducted in November 2014, in a randomized block design, with three replications. Populations and cultivars were evaluated in plots containing three lines with 6 m length, spaced at 0.7 m. One hundred and forty seeds were distributed in each plot, resulting in a seeding density of approximately 8 seeds per linear meter. The technical recommendations for cultivation in Brazil were adhered to consistently.

The traits evaluated at the individual level were grain yield (g per plant) and growth cycle (days). The growth cycle was defined as the days

between seedling emergence and full maturity (plants with 95% pods showing mature pod color) and the selection was practiced in both directions, i.e., for selection of early and late plants.

**Table 1.** Relationship of soybean segregant populations with their respective crosses and number of F2 individuals evaluated for grain yield, days to flowering and growth cycle days in Viçosa during the 2014/15 crop season.

Population	Parents	Number of F2 individuals evaluated
1	MSOY6001 × RSF6563IPRO	236
2	MSOY6001 × TMG123RR	265
3	MSOY6001 × SYN9078RR	197
4	MSOY6001 × TMG801	211
5	MSOY6001 × MSOY9144RR	262
6	RSF6563IPRO × TMG123RR	187
7	RSF6563IPRO × SYN9078RR	235
8	RSF6563IPRO × TMG801	226
9	RSF6563IPRO × MSOY9144RR	211
10	TMG123RR × SYN9078RR	209
11	TMG123RR × TMG801	240
12	TMG123RR × MSOY9144RR	221
13	SYN9078RR × TMG801	213
14	SYN9078RR × MSOY9144RR	188
15	TMG801 × MSOY144 RR	222
<b>Total</b>		<b>3323</b>

### Procedures for data analysis

Data were analyzed using REML/BLUP. The REML was applied to estimate the components of variance and BLUP was used to predict genotypic values. To apply the BLUPIS, the genotypic effect of each population was estimated by the following model:

$$y = X_r + Z_f + W_p + T_j + e \quad (1)$$

where  $y$  is the vector of phenotypic data;  $r$  is the vector of repetition effects (assumed as fixed) plus the general mean;  $f$  is the vector of population effects (assumed as random), in which,  $f \sim N(0, \sigma_f^2)$ ;  $p$  is the vector of plot effects (assumed as random, in which,  $p \sim N(0, \sigma_p^2)$ ;  $j$  is the vector of parents effects (assumed as random), in which,  $j \sim N(0, \sigma_j^2)$ ; and  $e$  is the vector of errors or residuals (random), in which,  $e \sim N(0, \sigma_e^2)$ . The capital letters (X, Z, W, and T) represent the incidence matrices for  $r, f, p$  and  $j$

effects, respectively. This model is based on the sum of the effects of population F2 plus the effects of individuals within populations. The model also provides the average genetic variability within population ( $\sigma_{ga}^2$ ) and the corresponding heritability ( $h_{adpop}^2$ ), estimated as described by Resende et al. (2016).

The BLUPIS estimates the number of individuals ( $n$ ) to be selected in each family ( $k$ ), based on the following equation proposed by Resende and Barbosa (2006):

$$n_k = \frac{\hat{g}_k}{\hat{g}_j} n_j \quad (2)$$

where  $\hat{g}_j$  refers to the genotypic effect of the best population,  $\hat{g}_k$  corresponds to the genotypic effect associated of  $k$ -th population, and  $n_j$  is equal to the number of individuals selected in the best population. To obtain the best performance of the BLUPIS method in comparison to BLUPI (Silva et al., 2015), 200 individuals were selected in the best population for each trait. The BLUPIS method automatically eliminates populations with a negative genotypic effect, i.e., those below the general mean of the experiment.

The genotypic value of each individual within the family, used in the BLUPI selection approach, was predicted by the adjustment of the following model:

$$y = Xr + Zg + Wp + e \quad (3)$$

where  $y$  is the vector of data;  $r$  is the vector of repetition effects (assumed to be fixed) added to the general mean;  $g$  is the vector of genotypic effects of populations plus individual effect within population (assumed as random), in which,  $g \sim N(0, \sigma_g^2)$ ;  $p$  is the vector of plot effects (random), in which,  $p \sim N(0, \sigma_p^2)$ ; and  $e$  is the vector of errors or residuals (random), in which,  $e \sim N(0, \sigma_e^2)$ . The capital letters (X, Z and W) represent the incidence matrices for  $r, g$  and  $p$  effects, respectively (Resende, 2007). This model requires an estimate of genetic variability within population ( $\sigma_{ga}^2$ ) and the corresponding heritability ( $h_{adpop}^2$ ) obtained in Eq. (1). Testing of model effects, Eq. (1) and Eq. (2), were conducted performing Likelihood Ratio Tests (LRT).

The selection of individuals by the BLUPI strategy was performed based on the classification of the genotypic value of all the individuals evaluated, according to the same number of genotypes indicated for selection using the BLUPIS procedure. Statistical analyzes were undertaken using Selegen-REML/BLUP software (Resende, 2016).

For the BLUPI method, the selection gain was obtained from the mean of the additive genetic values of the selected individuals. In this case, the selection was based on the population and individual information, with the ranking of the best individuals in the experiment equivalent to mass selection. In the case of BLUPIS, the selection of the best populations is considered first, with information inside, obtained through individual sampling. So, the main difference from BLUPI is the dynamic allocation of selection intensity within families/populations. The selection gain with the BLUPIS was estimated using the bootstrap resampling technique where the number of individuals to be selected in each population is sampled a thousand-fold at the 95% level of significance.

To compare the efficiency between the two methods (BLUPI and BLUPIS), a simple linear regression analysis (Silva et al., 2015) was performed, as follows:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (4)$$

where  $Y_i$  is the number of individuals selected by the BLUPI method in the  $i$  th population;  $X_i$  is the number of individuals selected for BLUPIS selection in the  $i$  th population;  $B_0$  is the intercept;  $B_1$  is the regression coefficient, and  $\varepsilon_i$  is the random error. Regression analyzes were performed using the Genes software (Cruz, 2013).

## Results

According to Likelihood Ratio Tests (LRT), there were effects of genotypes, plots and parents for growth cycle ( $p < 0.05$ ). For grain yield, the effects of genotypes ( $p < 0.10$ ) and plots ( $p < 0.05$ ) were important for the model fit. Therefore, the respective components of additive variance and heritability coefficients were statistically different from zero, which demonstrates the existence of genetic variability

and the possibility of gains with the selection for all the characters. In addition, the coefficient of inbreeding due to population differentiation ( $F_{ST}$ ) for grain yield was 0.75, on a scale of zero to one (Table 2). The  $F_{ST}$  provides the degree of differentiation or divergence between populations (Resende et al., 2016), so a high  $F_{ST}$  value indicates great genetic divergence.

**Table 2.** Estimates of variance components and genetic parameters for grain yield (g/plant) and growth cycle (days) for 15 soybean segregant populations in Viçosa, during the 2014/15 crop season.

Parameter <sup>1</sup>	Grain yield (g/plant)	Growth cycle (days)
$\sigma_{aF2}^2$	25.0436	41.5420
$\sigma_{plot}^2$	46.8270	4.4566
$\sigma_{ap}^2$	33.3078	132.3619
$\sigma_e^2$	1174.8720	68.1459
$\sigma_f^2$	1246.7426	114.1445
$\sigma_{fp}^2$	1255.0069	204.9644
$h_{aF2}^2$	0.0201	0.3639
$h_{ap}^2$	0.0265	0.6458
CVg (%)	8.8552	4.9556
CVe (%)	60.6522	6.3471
C <sub>2</sub> plot (%)	0.5500	1.8500
$F_{ST}$	0.7519	0.3139
$\sigma_{a0}^2$	16.6539	66.1809
$\sigma_{ad}^2$	4.1321	45.4099
$h_{adpop}^2$	0.0035	0.6664
$\sigma_{at}^2$	29.1757	86.9519
$h_{at}^2$	0.0234	0.7618
$h_{mp}^2$	0.7124	0.9458
$r_{\hat{a}ai}$	0.1417	0.6033
$r_{\hat{a}apop}$	0.9453	0.9976
$r_{\hat{a}aed}$	0.6604	0.8806
<b>General average</b>	<b>56.5131</b>	<b>130.0606</b>

<sup>1</sup>Additive variance ( $\sigma_{aF2}^2$ ); environmental variance between plots ( $\sigma_{plot}^2$ ); parents additive variance ( $\sigma_{ap}^2$ ); residual variance ( $\sigma_e^2$ ); individual phenotypic variance ( $\sigma_f^2$ ); phenotypic variance of the parents ( $\sigma_{fp}^2$ ); individual heritability in the restricted sense of F2 ( $h_{aF2}^2$ ); individual heritability in the restricted sense of the parents ( $h_{ap}^2$ ); coefficient of genetic variation among populations (CVg); coefficient of environmental variation (CVe); coefficient of determination of plot effects (C<sub>2</sub>plot); coefficient of inbreeding due to population differentiation ( $F_{ST}$ ); original additive genetic variance of the parents population ( $\sigma_{a0}^2$ ); mean genetic variance within population ( $\sigma_{ad}^2$ ); restricted mean heritability within population ( $h_{adpop}^2$ ); total genetic variance ( $\sigma_{at}^2$ ); total heritability ( $h_{at}^2$ ); mean heritability of population ( $h_{mp}^2$ ); accuracy of individual or mass selection ( $r_{\hat{a}ai}$ ); accuracy of population selection ( $r_{\hat{a}apop}$ ); accuracy of the combined selection between and within populations ( $r_{\hat{a}aed}$ ).

The coefficient of determination of the plot effects ( $C2_{plot}$ ) is a measure of experimental quality and quantifies the environmental variation between plots within blocks. The low  $C2_{plot}$  values (0.55 to 1.85%) observed, indicate low variability among plots within blocks and, consequently, the low environmental correlation between plants within the plot. Thus, these results are in accordance with the findings of Pimentel et al., (2014).

Table 2 presents the genetic parameters estimated for grain yield and growth cycle (days). In general, low estimates of restricted heritability of F2 between populations ( $h^2_{aF2}$ ) and restricted average heritability within populations ( $h^2_{adpop}$ ) were observed for the grain yield. This finding is reasonable, considering that these are F2 populations, where there is still a great

dominance effect and under such circumstances the coefficient associated with additive variance ( $\sigma_a^2$ ) is always equal to unity.

For the grain yield the combined selection, between and within by the BLUPIS method, determined the selection of 649 (19.53%) individuals distributed in the ten best populations (8, 6, 10, 15, 1, 9, 4, 7, 3 and 13), namely, those populations that presented a positive genotypic effect (Table 3). By the BLUPI method, the selection of the same number of individuals involved seven of the best populations (8, 6, 10, 15, 1, 9 and 4). Resende and Barbosa (2006) described the genotypic value, which corresponds to the sum of the genotypic effect and the general mean, as the best parameter to explain the superiority of a given cross. Therefore, it is expected that the best individuals are in these families.

**Table 3.** Values of genotypic effects ( $\hat{g}_j$ ), the genotypic value of individual population ( $\mu + \hat{g}_j$ ) in decreasing order, and the number of individuals selected using individual simulated BLUP (BLUPIS) and individual BLUP (BLUPI), for the grain yield, in segregating soybean populations.

Population	$\hat{g}_j$	$\mu + \hat{g}_j$	BLUPIS	BLUPI
8	5.5618	62.0749	200	226
6	5.5128	62.0259	198	187
10	2.2515	58.7646	81	209
15	1.1583	57.6714	42	18
1	1.1078	57.6209	40	6
9	1.0077	57.5207	36	2
4	0.6326	57.1457	23	1
7	0.4499	56.9630	16	0
3	0.2004	56.7135	7	0
13	0.1764	56.6895	6	0
12	-0.1325	56.3806	0	0
11	-0.5695	55.9436	0	0
14	-2.3762	54.1369	0	0
5	-6.5374	49.9757	0	0
2	-8.4435	48.0696	0	0
<b>Total</b>			<b>649</b>	<b>649</b>
<b>SG (%)</b>			<b>2.39</b> <b>[2.21; 2.57]<sup>1/</sup></b>	<b>8.73</b>

<sup>1/</sup>Confidence interval at the 95% probability level, estimated by the bootstrap resampling method, in the random selection, in each population, of the number of individuals established by the BLUPIS method, with one thousand repetitions. SG = Selection gain.

From the 15 studied populations for the growth cycle, eight (2, 1, 7, 6, 10, 11, 8 and 5) had negative genotypic effects and so favored the selection to reduce the number of days between emergence and complete maturation of the plants (Table 4). The BLUPIS strategy established the

selection of 19.47% (647) of the total number of individuals, distributed in eight populations while the selection of this same number of individuals by BLUPI involved only three populations. The selection of the 647 individuals in the three best populations initially seems to be advantageous

since it provides a high genetic gain of 5.70 versus 1.19% obtained with random selection via BLUPIS. However, this strategy also has the

lowest selective accuracy ever presented and low contribution to the maintenance of genetic variability in the breeding program.

**Table 4.** Values of genotypic effects ( $\hat{g}_j$ ), the genotypic value of individual population ( $\mu + \hat{g}_j$ ) in descending order, and the number of individuals selected using individual simulated BLUP (BLUPIS) and individual BLUP (BLUPI), for the growth cycle (days), separated into early and late, in segregating soybean populations.

Early growth cycle					Late growth cycle				
Population	$\hat{g}_j$	$\mu + \hat{g}_j$	BLUPIS	BLUPI	Population	$\hat{g}_j$	$\mu + \hat{g}_j$	BLUPIS	BLUPI
2	-8.9744	121.0862	200	262	15	17.5607	147.6213	200	222
1	-7.8287	122.2319	174	204	13	4.8312	134.8918	55	64
7	-7.6261	122.4345	170	181	14	4.3600	134.4206	50	38
6	-1.6286	128.4320	36	0	3	0.7483	130.8090	9	1
10	-1.2295	128.8311	27	0	4	0.7391	130.7997	8	0
11	-0.8498	129.2108	19	0	12	0.3948	130.4554	4	0
8	-0.6023	129.4583	10	0	9	0.3810	130.4416	4	0
5	-0.2757	129.7850	10	0	5	-0.2757	129.7850	0	1
9	0.3810	130.4416	0	0	8	-0.6023	129.4583	0	4
12	0.3948	130.4554	0	0	11	-0.8498	129.2108	0	0
4	0.7391	130.7997	0	0	10	-1.2295	128.8311	0	0
3	0.7483	130.8090	0	0	6	-1.6286	128.4320	0	0
14	4.3600	134.4206	0	0	7	-7.6261	122.4345	0	0
13	4.8312	134.8918	0	0	1	-7.8287	122.2319	0	0
15	17.5607	147.6213	0	0	2	-8.9744	121.0862	0	0
<b>Total</b>			<b>647</b>	<b>647</b>	<b>Total</b>			<b>330</b>	<b>330</b>
<b>SG (%)</b>			<b>1.19</b>	<b>5.7</b>	<b>SG (%)</b>			<b>3.49</b>	<b>11.39</b>
			[-0.88; 1.55] <sup>1/</sup>					[2.88; 4.06] <sup>1/</sup>	

<sup>1/</sup>Confidence interval at the 95% probability level, estimated by the bootstrap resampling method, in the random selection, in each population, of the number of individuals established by the BLUPIS method, with one thousand repetitions. SG = Selection gain.

In the selection for cycle extension (late cycle), 330 (9.93%) individuals were selected in seven populations (15, 13, 14, 3, 4, 12 and 9; Table 4). However, with the BLUPI-based selection, the same number of plants was selected from six of the fifteen populations, including two populations (5 and 8) that had a genotypic effect favorable for early maturation. The number of populations involved in the selection would indicate the contribution to the high genetic variability among the families to be advanced to the next generation. However, the advancement of these plants would bring great problems for the breeder, because he/she is also selecting populations that contribute to decreasing the average of the trait. In addition, segregation within families would require more labor to carry out a new stage of screening for the next generation cycle.

The BLUPIS and BLUPI methods were considered statistically equivalent if  $\hat{\beta}_0 = 0$  and

$\hat{\beta}_1 = 1$ , and the coefficient of determination ( $R^2$ ) of the model was higher than 70%. Table 5 lists the estimated regression parameters ( $\hat{\beta}_0$  and  $\hat{\beta}_1$ ), their respective statistical probabilities and the determination coefficients. These were statistically equal to zero and the unit, respectively, in four situations. For the grain yield, both parameters were different from zero and equal to one, respectively, proving the agreement between the two methods. In addition, the value of  $R^2$  (80.76%) indicated the model had a good fit in explaining the relation between the number of individuals selected by the two methods. In the other case, only the parameters  $\hat{\beta}_0$  for early cycle and late cycle were considered statistically equal to zero. Consequently, there was no optimal agreement between the methods for late and early cycle, even with the high levels of regression adjustment obtained.

**Table 5.** Estimates of constants ( $\hat{\beta}_0$ ) and coefficients of linear regression ( $\hat{\beta}_1$ ), p-values (P) with the hypotheses  $H_0: \hat{\beta}_0 = 0$  and  $H_0: \hat{\beta}_1 = 1$ , and determination coefficients ( $R^2$ ) associated with regressions between the total number of genotypes to be selected in each soybean population using individual BLUP (BLUPI), and the total number of genotypes to be selected in each population via individual simulated BLUP (BLUPIS) for the grain yield and growth cycle (early and late days).

Trait	$\hat{\beta}_0$	$\hat{\beta}_1$	$P(\hat{\beta}_0 = 0)$	$P(\hat{\beta}_1 = 1)$	$R^2$
Grain yield	-6.1264	1.1416	0.6245	0.6205	80.76
Early growth cycle	-9.7539	1.2280	0.0758	0.0026	96.87
Late growth cycle	-2.3618	1.1074	0.1715	1.53E-09	99.07

The correlation between the two methods was high for all the traits studied ( $r=0.8076 - 0.9907$ ). Nonetheless, the number of individuals selected per population in each technique presented different proportions, and so the regression showed a slope different from 45° or  $\hat{\beta}_1 \neq 1$ . Despite the lack of optimum agreement between the methods studied, the BLUPIS was the most indicated in the determination of the number of individuals to be advanced in a soybean breeding program.

## Discussion

The accuracy of individual selection (mass), considering the classification of Resende and Duarte (2007), was low for grain yield (0.14) and moderate for growth cycle (0.60), demonstrating that the possibility of gain with individual selection for grain yield is low (Table 2). Considering the selection of populations, the accuracy was very high for both characters studied (0.94~0.99) and, therefore, it allows to obtain greater gains with selection. For the combined selection, the selective accuracy between and within populations was moderate for grain yield (0.66) and high for growth cycle (0.88). These results show that the selection between and within populations provides greater gains than selection of the best individuals. The ratio of the accuracies of the selection responses between and within, and individuals gave the selection efficiencies as 4.66 and 1.46 for grain yield and growth cycle, respectively.

Of the total number of individuals selected by the BLUPI method, 622 (95.84%) belong to the three best populations (8, 6 and 10) while the remaining 5% were selected in the other four populations (Table 3). These results evidenced the low contribution of the BLUPI procedure to

the maintenance of genetic variability among the selected individuals, which is undesirable for a breeding program, which, when obtaining its elite lines, needs to recombine them for the continuity of the program.

In the BLUPIS procedure, more than 70% of the selected individuals were distributed among the three populations with superior genotypic effect relative to the others. These results suggest that the parents involved in these crosses had high specific combining ability, due to the complementarity provided by the presence of favorable alleles for the trait. Another relevant aspect is that these three populations share the cultivar RSF6563 IPRO as the parent, suggesting this parent has some general combinatorial ability and, as such, it has favorable alleles for the grain yield that are not found in the progenitors of the other populations.

The estimated gains from the selection of the two methods, BLUPIS (2.39%) and BLUPI (8.73%), were superior to those obtained by Barros et al. (2016) when practicing the selection in soybean genotypes in the F3 generation with different lateral branching abilities. However, they were lower than the gains documented by Reis et al. (2004). In these studies, the authors selected progenies in the F3 and F5 generations, which, because of their advancement for more generations of inbreeding, explored a larger fraction of the total additive variance (1.5~1.8) and, in turn, led to greater gains with the selection.

In the selection for the growth cycle, both early and late, the gains obtained with the BLUPI selection (5.7 and 11.39%) were higher than the BLUPIS gains (1.19 and 3.49%). However, it should be noted that in BLUPI selection for the early growth cycle, all selected individuals

belonged to the three best populations, and for the late growth cycle, this ratio was 98%. The superiority of the gains from BLUPI selection is therefore understandable, although these gains are difficult to reach due to the infeasibility of individual plant assessment in soybean breeding programs.

For the traits studied, the number of individuals selected per population decreased progressively from 200 in the population with the highest genetic effect, to zero in the population with no genetic effect. The results reflect the importance of the BLUPIS procedure in dynamically allocating the number of individuals selected per population, to the detriment of selection based on the individual genetic effect. According to the concept of effective population size of families S1 that is given by  $N_{ef} = [n/n + 0.5]$ , where  $N_{ef}$  maximum equals 1, when  $n$  tends to infinity, one can guarantee the selection of the best genotype with the selection of 50 plants within the best family ( $n_j = 50$ ) (Paternelli and Resende, 2015). It implies that 50 plants provide 99% of the maximum representativeness of an S1 family of full siblings and half-siblings.

Resende and Barbosa (2006) stated that BLUPIS is preferred for the genetic breeding of species whose data collection at the family level is operationally easier than at an individual level and so it is suitable for breeding programs of sugarcane, fodder species and annual autogamous species (soybeans, rice, beans, wheat and barley), especially for traits of low heritability. The efficiency of the BLUPIS method in the breeding of *Stylosanthes macrocephala* was validated by Resende et al. (2006). However, the authors emphasize that the method should be applied using the genotypic effects predicted via BLUP and not the genotypic values, much less the phenotypic mean of each family. In addition to misleading, determining

the number of individuals in the best family based on genotypic value and the phenotypic mean of each family would make the method similar to mass selection.

## Conclusion

The BLUPIS method presented superior performance when used in the selection of traits of quantitative genetic control.

The selection of the best populations, followed by the random selection of individuals within the populations was more efficient than the individual selection for all the characters studied.

BLUPIS procedure has the advantage of not selecting individuals in the families that performed below the general average. And BLUPIS contributes to the selection of individuals in representative numbers of the best populations, it can provide substantial gain for soybean breeding programs.

## Acknowledgements

We wish to thank Marcos Deon Vilela de Resende and Cosme Damião Cruz for implementing the routines and other genetic-statistical models free of charge for the entire scientific community through SELEGEN-REML\BLUP software ([http://www.ppestbio.ufv.br/?page\\_id=448](http://www.ppestbio.ufv.br/?page_id=448)) and GENES - a software package for analysis in experimental statistics and quantitative genetics (<http://arquivo.ufv.br/dbg/genes/genes.htm>), respectively. We appreciate and acknowledge the financial support from the National Council for Scientific and Technological Development (CNPq) and the Coordination for the Improvement of Higher Education Personnel (CAPES).

## References

- ALVES, R.S.; CARVALHO ROCHA, J.R.A.S.; TEODORO, P.E.; RESENDS, 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 Genetics & Genomes**, 14(77):1-8.



- BARROS, J.P.A.; SEDIYAMA, T.; SILVA, F.C.S.; SILVA, A.F.; BEZERRA, A.R.G.; ROSA, D.P.; LOPES, A.S.; OLIVEIRA, D.S. 2016. Estimates of genetic parameters and efficiency in selection for branching capacity in soybean genotypes. **Journal of Agronomy**, 15(1):39-44.
- CRUZ, C.D. 2013. GENES - A software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum Agronomy**, 35(3):271-276.
- GOWDA, M.; ZHAO, Y.; MAURER, H.P.; WEISSMANN, E.A.; WURSCHUM, T.; REIF, J.C. 2013. Best linear unbiased prediction of triticale hybrid performance. **Euphytica**, 191(1):223-230.
- LUZ, L.N.; SANTOS, R.; SILVA FILHO, J.L.; MELO FILHO, P.A. 2010. Estimativas de parâmetros genéticos em linhagens de amendoim baseadas em descritores associados ao ginóforo. **Revista Ciência Agrônômica**, 41(1):132-138.
- PEREIRA, T.B.; CARVALHO, J.P.F.; BOTELHO, C.E.; RESENDE, M.D.V.; REZENDE, J.C.; MENDES, A.N.G. 2013. Eficiência de seleção de progenies de café F<sub>4</sub> pela metodologia de modelos mistos (REML/BLUP). **Bragantia**, 72(3):230-236.
- PETERNELLI, L.A.; RESENDE, M.D.V. 2015. Experimental designs for next generation phenotyping. p. 15-32. In: FRITSCHÉ-NETO, R.; BORÉM, A.(ed.). Phenomics: How next-generation phenotyping is revolutionizing plant breeding. **Springer**, New York, USA.
- PIMENTEL, A.J.B.; GUIMARÃES, J.F.R.; SOUZA, M.A.; RESENDE, M.D.V.; MOURA, L.M.; CARVALHO, J.R.A.S.; RIBEIRO, G. 2014. Estimação de parâmetros genéticos e predição de valor genético aditivo de trigo utilizando modelos mistos. **Pesquisa Agropecuária Brasileira**, 49(11):882-890.
- REIS, E.F.; REIS, M.S.; CRUZ, C.D.; SEDIYAMA, T. 2004. Comparação de procedimentos de seleção para produção de grãos em populações de soja. **Ciência Rural**, 34(3):685-692.
- RESENDE, M.D.V. 2002. Genética biométrica e estatística no melhoramento de plantas perenes. Embrapa Informação Tecnológica, Brasília.
- RESENDE, M.D.V. 2007. SELEGEN-REML/BLUP: sistema estatístico e seleção genética computadorizada via modelos lineares mistos. Embrapa Florestas, Colombo.
- RESENDE, M.D.V. 2016. Software Selegen-REML/BLUP: a useful tool for plant breeding. **Crop Breeding and Applied Biotechnology**, 16(1):330-339.
- RESENDE, M.D.V.; BARBOSA, M.H.P. 2006. Selection via simulated BLUP based on family genotypic effects in sugarcane. **Pesquisa Agropecuária Brasileira**, 41(3):421-429.
- RESENDE, M.D.V.; DUARTE, J.B. 2007. Precisão e controle de qualidade em experimentos de avaliação de cultivares. **Pesquisa Agropecuária Tropical**, 37(3):182-194.
- RESENDE, R.M.S.; RESENDE, M.D.V.; LAURA, V.A.; JANK, L.; VALLE, C.B. 2006. Genotypic evaluation of accessions and individual selection in *Stylosanthes spp.* by simulated BLUP method. **Crop Breed Applied Biotechnology**, 6(1):253-260.
- RESENDE, M.D.V.; RAMALHO, M.A.P.; CARNEIRO, P.C.S.; CARNEIRO, J.E.S.; BATISTA, L.G.; GOIS, I.B. 2016. Selection index with parents, populations, progenies, and generations effects in autogamous plant breeding. **Crop Science**, 56(1):530-546.

- RESENDE, M.D.V.; ALVES, S.R. 2020. Linear, generalized, hierarchical, bayesian and random regression mixed models in genetics/genomics in plant breeding. **Functional Plant Breeding Journal**, 2(2):1-31.
- SILVA, F.L.; BARBOSA, M.H.P.; RESENDE, M.D.V.; PETERNELLI, L.A.; PEDROZO, C.A. 2015. Efficiency of selection within sugarcane families via simulated individual BLUP. **Crop Breeding and Applied Biotechnology**, 15(1):1-9.
- SPINELLI, V.M.; DIAS, L.A.S.; ROCHA, R.B.; RESENDE, M.D.V. 2015. Estimates of genetic parameters with selection within and between half-sib families of *Jatropha curcas* L. **Industrial Crops Products**, 69(1):355-361.
- VIANA, J.M.S.; MUNDIM, G.B.; LIMA, R.O.; SILVA, F.F.; RESENDE, M.D.V. 2014. Best linear unbiased prediction for genetic evaluation in reciprocal recurrent selection with popcorn populations. **Journal of Agricultural Science**, 152(1):4428-4438.