

## Analysis of direct and indirect selection and indices in soybean segregating populations

Marcelo Marchi Costa<sup>1\*</sup>, Antonio Orlando Di Mauro<sup>2</sup>, Sandra Helena Unêda-Trevisoli<sup>3</sup>, Nair Helena Castro Arriel<sup>4</sup>, Ivana Marino Bárbaro<sup>4</sup>, Gustavo Dias da Silveira<sup>6</sup>, and Franco Romero Silva Muniz<sup>2</sup>

Received 03 May 2007

Accepted 18 September 2007

**ABSTRACT** – *Estimates of gain with selection are very useful in breeding programs to predict the success of selection. Index-based simultaneous selection makes breeding more successful. The objective of this report was to estimate and compare the genetic gain obtained by direct and indirect selection and using the classical and based on desired gain indices. The experiment was set up in the design of families with intercalated checks with 293 F<sub>3</sub> soybean genotypes, distributed in 32 families derived from five crosses. Individual gains obtained with direct selection among and within families and mass selection were similar and in most cases higher than selection by indices. On the other hand, the highest total gains were obtained with selection indices and distributed across all traits. The classical index obtained the highest genetic gains.*

**Key words:** *Glycine max*, selection indices, expected gain, selection among and within families.

### INTRODUCTION

One of the main achievements of quantitative genetics is that it enables breeders to estimate the expected progress with selection before carrying it out. This information gives improvement programs a clearer orientation and helps predict the success of the adopted selection method and choose the technically most effective option on a scientific base (Ramalho et al. 1993, Cruz and Regazzi 1997).

The selection of superior progenies is a labor-intensive process, once the traits of importance are strongly influenced by the environment and often

correlated, so that a selection in one provokes effects in the others. Therefore, selection to develop superior genotypes based on only one or a few traits might be little effective, since a genotype may be obtained that performs superior in relation to the selected traits only (Cruz and Regazzi 1997, Cruz 2001).

The simultaneous selection of traits, which can be performed effectively by the use of selection indices, increases the chances for the success of breeding programs. The selection indices make a combination of the multiple information of the experimental unit possible and enable selection based on a complex of variables of economic interest (Cruz and Regazzi 1997).

<sup>1</sup> Departamento de Produção Vegetal, Faculdade de Ciências Agrárias e Veterinárias (FCAV), Universidade do Estado de São Paulo (UNESP), Campus de Jaboticabal. Via de Acesso Prof. Paulo Donato Castellane, s/nº, 14.884-900, Jaboticabal, SP, Brasil. \*Email: mmarchi@hotmail.com

<sup>2</sup> Departamento de Produção Vegetal, FCAV, UNESP

<sup>3</sup> APTA Regional Centro Leste, C.P. 271, 14.001-970, Ribeirão Preto, SP, Brasil

<sup>4</sup> Embrapa Algodão, C.P. 174, 58.107-720, Campina Grande, PB, Brasil

<sup>5</sup> APTA Regional Alta Mogiana, C.P. 35, 14.770-000, Colina, SP, Brasil

<sup>6</sup> Monsanto do Brasil, Rodovia BR-163, km 758, 78.890-000, Sorriso, MT, Brasil

Some comparisons of the indices with direct selection allow the conclusion that the use of indices as selection criterion achieves relatively superior results. In general, the direct gain for the trait is reduced; on the other hand, this reduction is compensated by a better distribution of favorable gains over the other traits. Different indices represent different options of selection and, consequently, of gains, which identify the most adequate genotypes for the breeder's objectives quickly and effectively (Cruz and Regazzi 1997). Several authors confirmed the efficacy of selection indices, among them Barbosa and Pinto (1998), Oliveira et al. (1999), Granate et al. (2002), Paula et al. (2002), and Costa et al. (2004).

This study aimed to estimate and compare the predicted genetic gains with direct and indirect selection, by the classical and the based on desired gains indices, for the selection of more promising soybean genotypes.

## MATERIAL AND METHODS

The experiment was conducted on an experimental area of the Faculdade de Ciências Agrárias e Veterinárias, UNESP – Campus de Jaboticabal. The genotypes were distributed on the field according to the family design with intercalated checks, with the control cultivars *Renascença* and *Liderança*. The technical recommendations for soybean cultivation were observed throughout the trial.

The plants of 293 genotypes of the  $F_3$  generation were used, representing 32 families derived from five crosses, and were evaluated in the  $R_8$  stage (Fehr and Caviness 1977) for the traits number of days to maturation (NDM); plant height at maturity (PHM) in cm; insertion height of the first pod (IHP) in cm; lodging (Lo), grade scale varying from 1 (upright/erect plant) to 5 (lodged plant); agronomic value (AV), grade scale varying from 1 (poor plant) to 5 (excellent plant); number of pods per plant (NP); number of seeds per plant (NS); and grain yield per plant (GY), in gram.

The variances of each trait for each control and for the segregating generation were analyzed according to the following statistical model:  $Y_{ij} = \mu + f_i + e_i + p_{ij} + \delta_{ij}$ , where  $Y_{ij}$  is an observation of the  $j^{\text{th}}$  plant of the  $i^{\text{th}}$  family;  $\mu$  is a general mean of the generation (control or family);  $f_i$  is the genetic effect attributed to the  $i^{\text{th}}$  family, with  $i = 1, 2, \dots, 32$ ;  $e_i$  is the environmental effect between rows (of the control or of families);  $p_{ij}$  is the genetic

effect ascribed to the  $j^{\text{th}}$  plant of the  $i^{\text{th}}$  family, with  $j = 1, 2, \dots, 293$ ;  $\delta_{ij}$  is the environmental effect between plants within rows (of the control or of families).

The statistical analyses were performed using software Genes (Cruz 2001). The original data of Lo and AV and of NP and NS were transformed in  $\sqrt{x+0.5}$  and  $\sqrt{x}$ , respectively, in order to adjust the data better to the normal distribution curve. The heritability coefficients were estimated by the parent-offspring regression, by the ratio of the covariance of parent ( $F_2$ ) and progeny ( $F_3$ ) generations by the variance of the parent generation, corrected by the inbreeding coefficient.

The selection gains were estimated by the following methods:

### Direct and indirect selection

The aim is to obtain selection gains in the target trait. The responses in the traits of secondary importance can be favorable or unfavorable and are not considered in the selection process, depending on the correlation between them.

### Selection among and within families

The expected gains among and within families through direct selection in trait  $x$  can be estimated by  $SG_x = i_e p h_x SD_{ex} + i_d p h_x SD_{dx}$ , where  $SG_x$  is the expected gain by direct selection among and within families in trait  $x$ ;  $i_e$  is the selection intensity obtained considering a selection of 50% among families;  $p$  is the parental control (considered equal to 1.0);  $h_x$  is the square root of the heritability of the parent-offspring regression of trait  $x$ ;  $SD_{ex}$  is the genetic standard deviation among families of trait  $x$ ;  $i_d$  is the selection intensity obtained considering a selection of 25% within families;  $SD_{dx}$  is the genetic standard deviation of trait  $x$  within families.

The indirect gain among and within families in trait  $y$  by selection for trait  $x$  ( $SG_{y(x)}$ ) is given by  $SG_{y(x)} = i_e p h_x r_e SD_{ey} + i_d p h_x r_d SD_{dy}$ , where  $r_e$  is the genetic correlation between traits  $x$  and  $y$ , among families;  $SD_{ey}$  is the genetic standard deviation among families of trait  $y$ ;  $r_d$  is the genetic correlation between traits  $x$  and  $y$ , within families;  $SD_{dy}$  is the genetic standard deviation within families for trait  $y$ .

### Mass selection

The expected gain is calculated based on the best progenies of the population by  $SG_x = i p h_x SD_{gx}$ , where  $SG_x$  is the direct gain expected by mass selection;  $i$  is

the selection intensity obtained considering a selection of 20% of the plants;  $h_x$  is the square root of the heritability of the parent-offspring regression of trait  $x$ ;  $SD_{gx}$  is the genetic standard deviation of trait  $x$ . In mass selection, indirect gains are also obtained by  $SG_{y(x)} = iph_x r_g SD_{gy}$ , where  $SG_{y(x)}$  is the gain obtained in trait  $y$  when selecting for trait  $x$ ;  $r_g$  is the genetic correlation between  $x$  and  $y$ ;  $SD_{gy}$  is the genetic standard deviation of trait  $y$ .

### Classical index of Smith (1936) and Hazel (1943)

The Classical index consists of the linear combination of several traits of economic importance whose weighting coefficients are estimated in order to maximize the correlation between the index and the genotypic aggregate. This aggregate is established by another linear combination, involving the genetic values, which are weighted for their respective economic values.

Let the selection index ( $I$ ) and the genotypic aggregate ( $H$ ) be described by  $I = b_1x_1 + b_2x_2 + \dots + b_nx_n = \sum b_ix_i = b'x$  and  $H = a_1g_1 + a_2g_2 + \dots + a_ng_n = \sum a_ig_i = a'g$ , where  $n$  is the number of evaluated traits;  $b'$  is the vector ( $1 \times n$ ) of the weighting coefficients of the selection index to be estimated;  $x$  is the matrix ( $n \times p$ ) of trait means;  $a'$  is the vector ( $1 \times n$ ) of previously established economic weights; and  $g$  is the matrix ( $n \times p$ ) of unknown genetic values of the  $n$  traits considered.

To estimate the selection index of each family the estimation of vector  $b$  is necessary, which is obtained so that the correlation between  $I$  and  $H$  is maximized. So,  $bP = Ga$ , where  $b$  is the estimator of vector ( $n \times 1$ ) of the weighting coefficients of the selection index;  $P$  is the matrix ( $n \times n$ ) of phenotypic variances and covariances of the traits; and  $G$  is the matrix ( $n \times n$ ) of genetic variances and covariances of the traits.

The expected gain for trait  $j$  in index-based selection is expressed by  $\Delta g_{j(I)} = b'G_j \frac{i}{\sqrt{V(I)}}$ , where  $\Delta g_{j(I)}$  is the gain for trait  $j$ , with selection based on index  $I$ ;  $b'$  is the vector of weighting coefficients of the traits in the selection index;  $G_j$  is the  $j^{\text{th}}$  row of matrix  $G$ ;  $i$  is the selection intensity for 20%; and  $V(I)$  is the index variance.

### Index based on the desired gains (Pesek and Baker, 1969)

In view of the difficulty of establishing the economic weights, an index was proposed where the

weights were replaced by the desired gains for each trait, of easier determination.

The construction of the index is based on the expected gain of the traits, given by  $\Delta g = \frac{Gbi}{\sigma_I}$ , where  $\Delta g$  is the gain estimated by the index;  $G$  is the matrix ( $n \times n$ ) of genetic variances and covariances among the traits;  $b$  is the vector ( $1 \times n$ ) of the weighting coefficients of the selection index to be estimated;  $i$  is the selection differential in units of standard deviation from index  $I$ ;  $\sigma_I$  is the standard deviation from index  $I$ . If  $\Delta g$  is replaced by  $\Delta g_d$ , which is the vector of the desired gains, and eliminating  $\frac{i}{\sigma_I}$ , which does not affect the proportion of the  $b$  coefficients,  $b$  can be estimated by the expression  $\hat{b} = G^{-1}\Delta g_d$ . The estimated  $b'_i$  coefficients allow the maximization of gains in each trait, according to the importance established by the breeder.

When using the classical index the genetic variation coefficient (CVg) and genetic standard deviation (SD) of each trait were considered as economic weights. SD was also used as desired gain with the Pesek and Baker index, as recommended by Cruz and Regazzi (1997). Furthermore, a negative selection for NDM and Lo was considered, to permit the development of earlier and less lodged genotypes. The traits AV, NP, NS, and GY were considered primary and NDM, PHM, IHP, and Lo secondary, with both indices.

Aiming at a standardized number of plants for each method, to obtain the estimates of gain prediction, selection intensities of 50 and 25% were adopted in the selection among and within families, respectively, and of 20% of the progenies for mass selection and the indices.

## RESULTS AND DISCUSSION

The estimates of the heritability coefficients based on parent-offspring regression are displayed in Table 1. Generally speaking, the highest coefficients were obtained for the traits NDM, NP, NS and GY, in most crosses. The values are determined by the high genotypic variance of these traits, due to the superiority of some plants and families in comparison with the population mean. These high heritabilities indicate a possible success in the selection of the early generations ( $F_3$ ) that were evaluated, in orientation of the selection process for the most promising genotypes.

**Table 1.** Estimates of the heritability coefficients based on the parent-offspring regression for eight evaluated traits in five soybean crosses

Crosses	NDM	PHM	IHP	Lo <sup>1</sup>	AV <sup>1</sup>	NP <sup>2</sup>	NS <sup>2</sup>	GY
1	0.81	0.76	0.39	0.30	0.40	0.82	0.65	0.41
2	0.72	0.68	0.41	0.55	0.59	0.81	0.91	0.88
3	0.47	0.86	0.81	0.86	0.60	0.92	0.92	0.76
4	0.93	0.77	0.75	0.90	0.81	0.74	0.88	0.84
5	0.87	0.69	0.53	0.73	0.83	0.72	0.70	0.86

<sup>1</sup> = data transformed by  $\sqrt{x+0.5}$ ; <sup>2</sup> = data transformed by  $\sqrt{x}$

\* NDM: number of days to maturity; PHM: plant height at maturity; IHP: insertion height of the first pod; Lo: lodging; AV: agronomic value; NP: number of pods; NS: number of seeds; GY: grain yield

### Estimates of selection gains

Tables 2 to 6 show the estimates of the gains with selection for the eight evaluated traits in the five crosses by direct and indirect selection and by the indices of Smith and Hazel (SH) and of Pesek and Baker (PB), using the previously described economic weights. Selection among and within families and mass were considered in the direct and indirect selection.

The highest direct gains were observed for the traits NP, NS and GY, in direct selection as well as by the indices (Tables 2 to 6). This fact had already been expected, since the selection was conducted prioritizing these traits with the highest heritability coefficients and high variation considered most important. Besides, it was stated that the individual gains obtained by direct selection among and within families and mass were similar, with a slight superiority of mass selection in the crosses 1 and 3 (Tables 2 and 4) for all traits, whereas selection among and within surpassed mass selection in the crosses 2, 4 and 5 (Tables 3, 5 and 6) in all traits as well, which were both superior to the SH and PB indices.

It must be emphasized that in some situations higher indirect than direct gains were observed in all five crosses, for the traits NP, NS and GY. This result is possible, according to Falconer (1987), if the heritability of the auxiliary trait is higher than of the main trait under selection and when the genetic correlation between both is of high magnitude.

Considering the total gains obtained by each selection criterion, a superiority of the SH index was observed, although with values close to those obtained by direct and indirect selection for the traits NP, NS and GY and the PB index in most situations, and with the exception of direct and indirect selection of NP in cross 1 (Table 2) and of the PB index in cross 2 (Table 3). The same fact was reported by Costa et al. (2004).

For trait NDM, where selection was performed for earliness, the direct gains obtained by selection among and within families and mass were very close. The former were superior in three of the five crosses (Tables 3, 5 and 6), and both outmatched the indirect gains and indices.

For PHM, the direct mass selection obtained slightly superior direct gains to those by selection among and within families in two crosses (Tables 2 and 4), with quite significant values in relation to those of indirect selection and by indices, except in some cases where the SH index obtained similar gains, especially with standard deviation as economic weight.

Likewise, the direct gains with mass selection exceeded selection among and within families in only two crosses for trait IHP (Tables 2 and 4), with marked superiority over the gains obtained by indirect selection and the indices, with exception of the SH index with standard deviation as weight in cross 3 (Table 4).

For the trait Lo, as in the case of NDM, selection was performed for the lowest values, to develop more erect plants. In this case, the direct gains by selection among and within families were a little superior to those of mass selection in three of the five crosses (Tables 3, 5 and 6) and both surpassed the indirect gains and indices by far.

An analysis of trait AV showed that the direct gains obtained by selection among and within families and mass were very close to the gains obtained by the indices, as well as to the indirect gains obtained by selection for NP, NS and GY in all crosses, which is most likely due to the high correlations between these traits. It was even observed that the indirect were higher than the direct gains (Table 2), which can be explained by the high correlation, as mentioned by Falconer (1987). The indirect gains were lower mainly in the selection for NDM, PHM, IHP, and Lo.

**Table 2.** Estimates of the selection gains (SG %) by the methods: direct and indirect selection, classical selection index of Smith & Hazel and selection index of Pesek & Baker, in the soybean cross 1 (MGBR 95-20937 x IAC Foscarin 31)

Trait	Method	SG (%)								
		NDM	PHM	IHP	Lo	AV	NP	NS	GY	Total
Direct and Indirect selection <sup>1</sup>										
NDM	ED <sup>2</sup>	<b>-4.49</b>	3.00	0.36	4.00	5.38	20.65	19.49	28.41	76.80
	M <sup>3</sup>	<b>-4.54</b>	3.18	1.51	4.32	5.74	24.22	22.54	34.00	90.97
PHM	ED	0.91	<b>13.79</b>	3.57	3.65	2.38	11.84	10.62	16.02	62.78
	M	0.97	<b>13.93</b>	3.85	3.55	2.36	12.21	10.93	15.65	63.45
IHP	ED	0.04	1.46	<b>17.50</b>	0.34	0.46	1.76	1.14	-0.09	22.61
	M	0.19	1.58	<b>17.67</b>	0.29	0.47	3.51	2.50	3.15	29.36
Lo	ED	1.10	3.31	0.75	<b>-6.00</b>	0.66	5.40	4.99	7.54	17.75
	M	1.19	3.22	0.65	<b>-6.06</b>	0.22	6.03	5.33	8.73	19.31
AV	ED	1.84	2.69	1.28	0.83	<b>6.43</b>	21.54	20.30	28.22	83.13
	M	1.95	2.67	1.28	0.27	<b>6.50</b>	21.63	20.41	28.66	83.37
NP	ED	2.72	5.14	1.85	2.58	8.28	<b>34.44</b>	32.17	44.44	131.62
	M	3.19	5.31	3.72	2.89	8.31	<b>34.80</b>	32.51	45.12	135.85
NS	ED	2.42	4.35	1.14	2.25	7.36	30.36	<b>28.62</b>	39.55	116.05
	M	2.80	4.48	2.49	2.41	7.40	30.68	<b>28.92</b>	39.79	118.97
GY	ED	2.00	3.71	-0.05	1.92	5.78	23.69	22.34	<b>32.47</b>	91.86
	M	2.39	3.63	1.78	2.23	5.87	24.05	22.48	<b>32.81</b>	95.24
	SH - CVg <sup>4</sup>	6.51	12.88	5.62	2.01	6.49	33.16	27.98	30.43	125.08
	SH - SD <sup>5</sup>	4.89	13.02	2.23	2.34	3.85	28.54	20.63	23.92	99.42
	Pesek & Baker	4.39	7.99	0.68	1.16	6.36	31.38	26.82	32.35	111.13

<sup>1</sup> Values in bold correspond to the direct selection gains for the target trait <sup>2</sup> ED: selection among and within families. <sup>3</sup> M: mass selection <sup>4</sup> SH - CVg: selection by the classical index, using the genetic variation coefficient (CVg) as economic weight. <sup>5</sup> SH - SD: selection by the classical index, genetic standard deviation (SD) as economic weight \* NDM: number of days to maturity; PHM: plant height at maturity; IHP: insertion height of the first pod; Lo: lodging; AV: agronomic value; NP: number of pods; NS: number of seeds; GY: grain yield

**Table 3.** Estimates of selection gains (SG %) by the methods: direct and indirect selection, classical selection index of Smith & Hazel and Pesek & Baker selection index, in the soybean cross 2 (MGBR 95-20937 x BR-16)

Trait	Method	SG (%)								
		NDM	PHM	IHP	Lo	AV	NP	NS	GY	Total
Direct and Indirect selection <sup>1</sup>										
NDM	ED <sup>2</sup>	<b>-2.78</b>	1.20	-2.18	2.66	-1.16	0.28	-0.39	-3.18	-5.55
	M <sup>3</sup>	<b>-2.70</b>	1.18	-2.09	2.21	-1.08	0.11	-0.63	-3.23	-6.23
PHM	ED	0.27	<b>11.85</b>	6.88	2.63	-0.46	1.82	3.83	3.18	30.00
	M	0.26	<b>11.50</b>	6.74	2.46	-0.69	0.20	1.80	1.27	23.54
IHP	ED	-0.24	3.48	<b>14.28</b>	0.25	-2.80	-8.56	-8.17	-13.50	-15.26
	M	-0.23	3.41	<b>13.86</b>	0.28	-2.68	-8.42	-8.12	-13.10	-15.00
Lo	ED	0.87	3.86	0.73	<b>-6.54</b>	-2.84	-1.60	-1.95	-6.65	-14.12
	M	0.72	3.63	0.82	<b>-6.35</b>	-2.68	-1.48	-1.80	-5.97	-13.11
AV	ED	-0.30	-0.54	-6.55	-2.27	<b>8.72</b>	19.66	21.77	36.05	76.54
	M	-0.28	-0.81	-6.27	-2.14	<b>8.46</b>	19.16	21.15	35.23	74.50
NP	ED	0.03	0.92	-8.53	-0.55	8.39	<b>27.98</b>	30.59	50.10	108.93
	M	0.01	0.10	-8.40	-0.50	8.17	<b>27.16</b>	29.76	48.66	104.96
NS	ED	-0.04	1.82	-7.69	-0.63	8.77	28.89	<b>33.58</b>	52.95	117.65
	M	-0.07	0.86	-7.64	-0.58	8.52	28.10	<b>32.60</b>	52.40	114.19
GY	ED	-0.19	0.88	-7.35	-1.24	8.40	27.37	31.21	<b>55.90</b>	114.98
	M	-0.20	0.35	-7.13	-1.11	8.21	26.58	30.30	<b>54.26</b>	111.26
	SH - CVg <sup>4</sup>	-0.06	3.46	-2.38	-1.74	8.23	26.60	32.85	52.86	119.82
	SH - SD <sup>5</sup>	-0.09	4.67	-1.11	-1.50	7.98	26.03	32.26	51.65	119.89
	Pesek & Baker	-0.03	1.64	-4.79	-1.18	8.71	27.12	33.12	54.90	119.49

<sup>1</sup> Values in bold correspond to the direct selection gains for the target trait. <sup>2</sup> ED: selection among and within families. <sup>3</sup> M: mass selection. <sup>4</sup> SH - CVg: selection by the classical index, using the genetic variation coefficient (CVg) as economic weight. <sup>5</sup> SH - SD: selection by the classical index, genetic standard deviation (SD) as economic weight. \* NDM: number of days to maturity; PHM: plant height at maturity; IHP: insertion height of the first pod; Lo: lodging; AV: agronomic value; NP: number of pods; NS: number of seeds; GY: grain yield

**Table 4.** Estimates of the selection gains (SG %) by the methods: direct and indirect selection, classical selection index of Smith & Hazel and Pesek & Baker selection index, in the soybean cross 3 (Renasçença x IAC Foscarin 31)

Trait	Method	SG (%)								
		NDM	PHM	IHP	Lo	AV	NP	NS	GY	Total
Direct and Indirect selection <sup>1</sup>										
NDM	ED <sup>2</sup>	<b>-2.58</b>	-5.77	-7.67	-4.38	1.36	9.41	8.73	12.16	11.26
	M <sup>3</sup>	<b>-2.66</b>	-5.49	-7.08	-4.16	0.68	8.85	7.70	8.83	6.67
PHM	ED	-1.57	<b>17.22</b>	10.44	7.86	0.12	-5.33	-4.93	-4.68	19.13
	M	-1.50	<b>17.74</b>	11.46	8.04	0.11	-8.64	-7.35	-5.58	14.28
IHP	ED	-1.22	6.10	<b>27.78</b>	-4.05	0.08	-7.87	-7.59	-7.85	5.38
	M	-1.13	6.69	<b>28.61</b>	-4.16	0.17	-9.43	-8.63	-7.71	4.41
Lo	ED	-1.23	8.09	-7.14	<b>-16.76</b>	-3.79	-5.75	-5.91	-10.41	-42.90
	M	-1.17	8.27	-7.33	<b>-17.27</b>	-3.69	-6.36	-6.14	-9.56	-43.25
AV	ED	0.52	0.16	0.19	-5.14	<b>8.55</b>	21.72	21.97	35.04	83.01
	M	0.26	0.15	0.41	-5.00	<b>8.80</b>	22.12	22.47	35.84	85.05
NP	ED	1.40	-2.89	-7.31	-3.03	8.44	<b>33.99</b>	33.38	49.78	113.76
	M	1.31	-4.68	-8.75	-3.53	8.59	<b>35.01</b>	34.38	50.92	113.25
NS	ED	1.30	-2.69	-7.08	-3.13	8.58	33.57	<b>33.54</b>	50.41	114.50
	M	1.14	-4.01	-8.05	-3.25	-8.78	34.58	<b>34.55</b>	51.72	97.90
GY	ED	1.07	-1.51	-4.33	-3.26	8.03	29.54	29.74	<b>47.34</b>	106.62
	M	0.77	-1.79	-4.25	-2.99	8.26	30.21	30.52	<b>48.76</b>	109.49
	SH – CVg <sup>4</sup>	0.13	3.21	13.50	-4.55	8.08	31.78	31.24	41.64	125.03
	SH – SD <sup>5</sup>	-0.73	14.78	28.05	1.69	5.66	17.03	18.18	37.35	122.01
	Pesek & Baker	0.80	0.15	-7.91	-2.93	8.60	33.97	33.63	47.84	114.15

<sup>1</sup> Values in bold correspond to the direct selection gains for the target trait. <sup>2</sup> ED: selection among and within families. <sup>3</sup> M: mass selection. <sup>4</sup> SH – CVg: selection by the classical index, using the genetic variation coefficient (CVg) as economic weight. <sup>5</sup> SH – SD: selection by classical index, genetic standard deviation (SD) as economic weight. \* NDM: number of days to maturity; PHM: plant height at maturity; IHP: insertion height of the first pod; Lo: lodging; AV: agronomic value; NP: number of pods; NS: number of seeds; GY: grain yield

For trait NP, the direct gains obtained by mass selection and among and within families was also very close to those of the indices SH and PB and to the indirect gains for AV, NS and GY, also owing to the high correlations. The presence of some indirect that surpassed the direct gains (Tables 3, 5 and 6) was also observed, based on the explanations of Falconer (1987) as well, and indirect gains of considerable magnitude in the selection for NDM in cross 1 (Table 2). With exception of the SH index with standard deviation as weight in cross 3 (Table 4), the gains obtained with the use of the indices were close to those of direct selection throughout.

As for NP, the direct gains obtained by selection among and within families and mass for trait NS were close to the indirect gains with selection for AV, NP and GY, due to their high correlations and to the gains obtained with the indices, except for the SH index with standard deviation as economic weight in cross 3 (Table 4). The presence of higher indirect than direct gains was also observed (Tables 2 and 6), and considerable values in the selection for NDM in cross 1 (Table 2).

When analyzing trait GY, the most important from the economic point of view, mass selection resulted in

higher direct gains than those of selection among and within families in two crosses (Tables 2 and 4). These values were close, mainly, to the indirect gains obtained with selection for AV, NP and NS, and even achieved higher indirect gains in some cases (Tables 2, 4 and 5), which can be explained by the high correlations and heritabilities (Falconer 1987). Furthermore, the high values of indirect gains in the selection for NDM in cross 1 (Table 2) are worth mentioning. The gains obtained by the indices were quite close to the direct gains, with a slight inferiority of the SH index with the standard deviation as weight in the crosses 1, 3 and 4 (Table 2, 4 and 5).

It is noteworthy that the closeness of the gains obtained with direct selection, by the mass method as much as among and within families, to those obtained by the SH and PB indices, especially for the traits AV, NP, NS and GY, is very likely due to the fact that these traits were considered primary in the estimates of the indices, so the selection was mainly based on these traits.

In a comparison of our results to those of Oliveira et al. (1999), the direct gains of this experiment were similar for NP and NS and inferior for GY. The indirect gains for

**Table 5.** Estimates of the selection gains (SG %) by the methods: direct and indirect selection, classical selection index of Smith & Hazel and Pesek & Baker selection index, in the soybean cross 4 (Renasçença x IAC-17)

Trait	Method	SG (%)								
		NDM	PHM	IHP	Lo	AV	NP	NS	GY	Total
Direct and Indirect selection <sup>1</sup>										
NDM	ED <sup>2</sup>	<b>-3.64</b>	5.31	2.19	-4.51	5.86	12.95	14.13	20.77	53.06
	M <sup>3</sup>	<b>-3.56</b>	4.09	3.03	-2.95	5.03	11.06	9.80	15.48	41.98
PHM	ED	2.15	<b>7.49</b>	2.06	-0.75	5.05	10.67	11.35	17.75	55.77
	M	1.66	<b>7.33</b>	2.87	3.00	3.18	6.82	5.72	9.51	40.09
IHP	ED	0.21	0.47	<b>30.97</b>	-0.31	-3.31	-7.44	-7.94	-13.04	-0.39
	M	0.29	0.68	<b>30.31</b>	-0.38	-2.97	-6.75	-5.86	-11.66	3.66
Lo	ED	-1.68	-0.69	-1.19	<b>-9.49</b>	-6.50	-11.53	-12.52	-21.29	-64.89
	M	-1.10	2.76	-1.48	<b>-9.29</b>	-5.16	-8.21	-7.21	-14.52	-44.21
AV	ED	1.65	3.50	-9.72	-4.89	<b>11.30</b>	21.45	23.26	36.75	83.30
	M	1.41	2.20	-8.74	-3.89	<b>11.06</b>	20.75	18.34	34.89	76.02
NP	ED	1.71	3.48	-10.28	-4.09	10.10	<b>21.91</b>	23.68	37.68	84.19
	M	1.46	2.22	-9.32	-2.91	9.77	<b>21.44</b>	18.89	36.53	78.08
NS	ED	1.87	3.70	-11.00	-4.44	10.95	23.68	<b>26.05</b>	41.13	91.94
	M	1.59	2.28	-9.93	-3.13	10.57	23.14	<b>20.82</b>	39.91	85.25
GY	ED	1.66	3.50	-10.91	3.14	10.48	22.82	23.91	<b>40.97</b>	95.57
	M	1.24	1.88	-9.76	-3.12	9.94	22.13	19.74	<b>40.10</b>	82.15
	SH - CVg <sup>4</sup>	1.80	4.68	2.16	-4.40	10.94	19.96	21.63	39.47	96.24
	SH - SD <sup>5</sup>	1.96	7.21	3.49	-1.10	11.22	17.21	20.87	35.33	96.19
	Pesek & Baker	1.78	3.97	-9.19	-5.96	11.19	21.13	26.00	40.57	89.49

<sup>1</sup> Values in bold correspond to the direct selection gains for the target trait. <sup>2</sup> ED: selection among and within families. <sup>3</sup> M: mass selection.

<sup>4</sup> SH - CVg: selection by the classical index, using the genetic variation coefficient (CVg) as economic weight. <sup>5</sup> SH - SD: selection by the classical index, genetic standard deviation (SD) as economic weight. \* NDM: number of days to maturity; PHM: plant height at maturity; IHP: insertion height of the first pod; Lo: lodging; AV: agronomic value; NP: number of pods; NS: number of seeds; GY: grain yield

NP obtained here were similar to the selection for NS and inferior to the selection for GY, as much as with NS, with similar gains in the selection for NP and lower ones in the selection for GY. The indirect gains for GY with selection for NP and NS obtained here were lower though, and the gains by the PB index were higher for NP, similar for NS and lower for GY.

Similar amplitudes of gains were obtained by Reis et al. (2002), in the selection among and within families as much as in the mass selection for NDM, PHM and NP. For the latter two, the authors obtained some superior gains. Lower gains were observed by Reis et al. (2002) regarding trait GY.

The indirect gains for NDM and PHM, as well as the direct and indirect gains for NP and GY, were lower than the ones obtained by Backes et al. (2003) with selection for NP and GY, in most situations. With regard to the gains obtained by the Pesek and Baker index, the results of this experiment were somewhat inferior to those obtained by Backes et al. (2003) for NP and GY, and quite different for PHM, once the authors performed negative selection for this trait.

The direct gains for IHP, NP, NS and GY were moreover similar to those obtained by Costa et al. (2004), but inferior for PHM, Lo and AV. The indirect gains obtained here by selection for the other traits were similar for PHM, IHP, and Lo, while for AV, NP and NS some values were similar and others lower, and for GY some were similar and others higher. In respect of the gains obtained by the SH index, the results observed here were similar for all traits, while with the PB index, the gains for NP and NS observed by Costa et al. (2004) were lower, but similar for PHM, IHP, Lo, AV, and GY.

Based on our results of estimated gains, the crosses 3, based on the total gains, and 5, considering direct gains for GY, appear to be the most promising crosses. Cross 4 was the least promising for total gains and cross 1 for direct gains in GY.

The results obtained by the different methods, despite little differentiated, allow the inference that the use of the indices is advantageous over direct selection, since the obtained gains are distributed among all evaluated traits and achieve a higher total,

**Table 6.** Estimates of the selection gains (SG %) by the methods: direct and indirect selection, selection index of Smith & Hazel and Pesek & Baker selection index, in the soybean cross 5 (Liderança x IAC-17)

Trait	Method	SG (%)								
		NDM	PHM	IHP	Lo	AV	NP	NS	GY	Total
Direct and Indirect selection <sup>1</sup>										
NDM	ED <sup>2</sup>	<b>-4.48</b>	9.04	-0.16	-0.15	3.49	6.18	4.86	6.98	25.76
	M <sup>3</sup>	<b>-4.33</b>	8.83	-0.04	0.91	3.39	6.04	4.80	7.10	26.70
PHM	ED	2.60	<b>12.37</b>	-4.58	3.88	4.93	11.09	9.36	16.93	56.58
	M	2.54	<b>11.96</b>	-3.75	3.48	4.67	11.60	9.31	18.03	57.84
IHP	ED	-0.03	-3.27	<b>13.40</b>	-1.91	-3.25	-10.87	-9.36	-21.30	-36.59
	M	-0.01	-2.67	<b>12.96</b>	-0.76	-2.98	-11.03	-9.10	-21.36	-34.95
Lo	ED	-0.06	5.78	-3.99	<b>-8.84</b>	1.23	7.19	5.44	10.47	17.22
	M	0.39	5.19	-1.58	<b>-8.55</b>	0.40	5.36	3.56	6.58	11.35
AV	ED	1.32	6.49	-6.00	1.09	<b>11.38</b>	24.39	21.37	49.55	109.59
	M	1.28	6.15	-5.51	0.36	<b>11.01</b>	23.64	20.81	47.85	105.59
NP	ED	0.90	5.56	-7.63	2.42	9.28	<b>25.90</b>	22.62	53.76	112.81
	M	0.87	5.81	-7.75	1.80	9.00	<b>25.06</b>	21.91	52.17	108.87
NS	ED	0.76	5.18	-7.25	2.02	8.97	24.96	<b>22.83</b>	53.68	111.15
	M	0.76	5.15	-7.06	1.32	8.74	24.18	<b>22.09</b>	51.79	106.97
GY	ED	0.50	4.22	-7.43	1.75	9.37	26.72	24.17	<b>59.82</b>	119.12
	M	0.51	4.49	-7.46	1.10	9.05	25.93	23.32	<b>59.80</b>	116.74
	SH - CVg <sup>4</sup>	0.90	5.70	-7.26	1.91	11.12	24.97	22.36	59.84	119.54
	SH - SD <sup>5</sup>	1.03	6.58	-7.73	2.33	11.01	24.92	22.17	59.30	119.61
	Pesek & Baker	0.77	2.64	-5.33	-0.61	11.38	22.89	22.40	59.84	113.98

<sup>1</sup> Values in bold correspond to the direct selection gains for the target trait. <sup>2</sup> ED: selection among and within families. <sup>3</sup> M: mass selection  
<sup>4</sup> SH - CVg: selection by the classical index, using the genetic variation coefficient (CVg) as economic weight. <sup>5</sup> SH - SD: selection by the classical index, genetic standard deviation (SD) as economic weight. \* NDM: number of days to maturity; PHM: plant height at maturity; IHP: insertion height of the first pod; Lo: lodging; AV: agronomic value; NP: number of pods; NS: number of seeds; GY: grain yield

without a significant loss in the main trait, giving the classical index of Smith and Hazel a slight advantage. The efficacy of the use of the selection indices was observed by Barbosa and Pinto (1998), Oliveira et al. (1999), Granate et al. (2002), Paula et al. (2002), as well as Costa et al. (2004).

#### ACKNOWLEDGEMENTS

The Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) supported this report with funds, and the technical assistants Amanda Roberta Corrado and Geraldo Mangela de Assis helped conduct the experiment.

## Análise da seleção direta, indireta e por índices em populações segregantes de soja

**RESUMO** - As estimativas de ganho com a seleção auxiliam no direcionamento dos programas de melhoramento e predição de sucesso no processo seletivo, sendo a seleção simultânea, através de índices de seleção, responsável pelo incremento no êxito do melhorista. Assim, o objetivo do presente trabalho foi estimar e comparar os ganhos genéticos obtidos pela seleção direta e indireta e pelos índices clássico e baseado nos ganhos desejados. Utilizou-se o delineamento de famílias com testemunhas intercalares, com 293 genótipos da geração F<sub>3</sub>, distribuídos em 32 famílias oriundas de cinco cruzamentos. Os ganhos individuais pela seleção direta entre e dentro de famílias e massal foram próximos e superiores aos índices, na maioria das situações. Por sua vez, os maiores ganhos totais ocorreram nos índices de seleção, os quais foram distribuídos entre todos os caracteres, destacando-se o índice clássico com os maiores ganhos genéticos.

**Palavras-chave:** *Glycine max*, índices de seleção, ganho esperado, seleção entre e dentro de famílias.



## REFERENCES

- Backes RL, Reis MS, Cruz CD, Sedyama T and Sedyama CS (2003) Correlation estimates and assessment of selection strategies in five soybean populations. **Crop Breeding and Applied Biotechnology** 3: 107-116.
- Barbosa MHP and Pinto CABP (1998) Eficiência de índices de seleção na identificação de clones superiores de batata. **Pesquisa Agropecuária Brasileira** 33: 149-156.
- Costa MM, Mauro AO, Unêda-Trevisoli SH, Arriel NHC, Bárbaro IM and Muniz FRS (2004) Ganho genético por diferentes critérios de seleção em populações segregantes de soja. **Pesquisa Agropecuária Brasileira** 39: 1095-1102.
- Cruz CD and Regazzi AJ (1997) **Modelos biométricos aplicados ao melhoramento genético**. 2<sup>nd</sup> ed., Editora UFV, Viçosa, 390p.
- Cruz CD (2001) **Programa genes: versão Windows; aplicativo computacional em genética e estatística**. Editora UFV, Viçosa, 648p.
- Falconer DS (1987) **Introdução à genética quantitativa**. Tradução de Martinho de Almeida e Silva e José Carlos Silva. UFV-Imprensa Universitária, Viçosa, 279p.
- Fehr WR and Caviness JA (1977) **Stages of soybean development**. Iowa State University, Ames, 11p. (Special Report 80)
- Granate MJ, Cruz CD and Pacheco CAP (2002) Predição de ganho genético com diferentes índices de seleção no milho pipoca CMS-43. **Pesquisa Agropecuária Brasileira** 37: 1001-1008.
- Hazel LN (1943) The genetic basis for constructing selection indexes. **Genetics** 28: 476-490.
- Oliveira ACB, Sedyama CS and Cruz CD (1999) Selection for later flowering in soybean (*Glycine max* L. Merrill) F<sub>2</sub> populations cultivated under short day conditions. **Genetics and Molecular Biology** 22: 243-247.
- Paula RC, Pires IE, Borges RCG and Cruz CD (2002) Predição de ganhos genéticos em melhoramento florestal. **Pesquisa Agropecuária Brasileira** 37: 159-165.
- Pesek J and Baker RJ (1969) Desired improvement in relation to selected indices. **Canadian Journal of Plant Science** 49: 803-804.
- Ramalho MAP, Santos JB and Zimmermann MJO (1993) **Genética quantitativa em plantas autógamas: aplicações ao melhoramento do feijoeiro**. Editora UFG, Goiânia, 271p.
- Reis EF, Reis MS, Cruz CD, Sedyama T and Teixeira RC (2002) Predicted and realized gains of different selection strategies in soybean populations: a comparative study. **Crop Breeding and Applied Biotechnology** 2: 61-70.
- Smith HF (1936) A discriminant function for plant selection. **Annals of Eugenics** 7: 240-250.