

Genotype x environment interactions and implications for associations among soybean traits

D.R. Pereira¹, A.T. Bruzi¹, J.A.R. Nunes², J.P.S. Carvalho¹, E.V. Zambiazzi³ and G.S. Gesteira⁴

¹ Departamento de Agricultura, Universidade Federal de Lavras, Lavras, MG, Brasil

² Departamento de Biologia, Universidade Federal de Lavras, Lavras, MG, Brasil

³ Bayer, Cristalina, GO, Brasil

⁴ Departamento de Genética – LGN, Escola Superior de Agricultura Luiz de Queiroz - Universidade de São Paulo, Piracicaba, SP, Brasil

Corresponding author: D.R. Pereira
E-mail: dyannarangel@hotmail.com

Genet. Mol. Res. 18 (3): gmr18099
Received August 01, 2018
Accepted January 25, 2019
Published September 25, 2019
DOI <http://dx.doi.org/10.4238/gmr18099>

ABSTRACT. An option into guide selection strategies in soybean breeding programs is to investigate associations between the main traits studied by breeders. A question that is faced is the possible influence of genotype x environment (GE) interactions on correlations among traits. Path analysis allows, in addition to measuring correlation, a more detailed study of cause / effect relationships among traits. We measured the effect of the GE interaction in the association between agronomic traits in soybeans. The experiments were carried out during two crop seasons, in three municipalities of Minas Gerais, Brazil, using a randomized complete blocks design, with 35 treatments. We evaluated number of grains, number of pods (NP), number of grains per pod, grain yield, weight of 100 grains, lodging score, plant height, height of first pod insertion, and full maturity. The data were submitted to variance analysis and path analysis. We also calculated the sum of ranks index in three ways, considering different economic weights. The correlation between grain yield and other traits, as well as direct and indirect effects, varied in magnitude and direction, considering individual and multi-environmental analyses, demonstrating the influence of the GE interaction. For example, considering the individual analysis for Lavras

in 2015/16, the correlation between grain yield and NP was -0.0534 and was not significant at the 5% level by the Mantel test, and the direct effect of NP on grain yield was 0.1717, while in the joint analysis of all environments, the correlation between these traits was 0.3674 and was significant, and the direct effect of NP on grain yield was -0.0520. The direct effect as an economic weight in the sum of rank index can be more efficient in selecting cultivars when compared to a simple correlation; this could become a useful strategy for the selection of multiple traits.

Key words: *Glycine max*; Grain yield; Plant breeding; Path analysis

INTRODUCTION

Despite the great utility of simple correlation coefficient estimates for the understanding of a complex trait, such as grain yield, they are not cause-effect measures, since this quantification is restricted to two-by-two comparisons and does not consider direct and indirect influences of other traits (Neto et al., 2011; Nogueira et al., 2012). Thus, if a high correlation between two traits is a consequence of the indirect effect of another one, the estimation may be misleading, resulting in mistaken selection strategies (Perini et al., 2012).

Therefore, it is useful and necessary for the breeder to detail the degree of relationship of the traits through an appropriate method, such as path analysis. This method allows unfolding the estimated correlation coefficients of the explanatory component traits into their direct and indirect effects on most traits (Teodoro et al., 2015), helping to understand the real relationships between a set of characters and a main variable (Baraskaret al., 2015).

However, the genotype x environment (GE) interaction may greatly influence complex traits, such as grain yield, which can be explained by the effect of various characters on yield components. Thus, GE interaction may exert an important influence on the association among traits, which can affect the selection process, especially when it involves indirect selection and multiple traits selection, which in the latter case interferes with the recommendation of adapted cultivars for different environments (Lopes et al., 2002).

Moreover, the main objective of breeding programs is to obtain cultivars that accumulate advantages and that are capable of surpassing those already in the market, i.e., cultivars that simultaneously provide favorable phenotypes for several traits. Thus, selection based on one or few characteristics becomes inadequate, since it leads to a superior product only in relation to the characters by which it was selected (Cruz and Carneiro, 2014).

An alternative to the selection of multiple characters is index selection, which has been reported in soybean crop breeding (Soares et al., 2015; Bizari et al., 2017).

The breeders' preference is generally for indexes of easy application and interpretation (Reis et al., 2015), such as the sum of ranks index. This index allows combining various types of information of the evaluated traits that are chosen by the researcher based on the simultaneous selection desired, obtaining a single value that involves all the others and works as if it were an additional character (Cruz and Carneiro, 2014).

However, a pending issue concerns the difficulty of specifying the economic weights for each character in the selection index (Santos et al., 2007). Usually, breeders do not have these economic weights fully determined. Consequently, proposals with varied alternative weights are common. For Resende (2007), the use of weights calculated based on the correlation between explanatory characters and the basic character of interest is more appropriate.

Following similar reasoning, setting aside the simple correlation unfolding by path analysis, another possibility is to calculate these weights as a function of the direct effects.

Along this line, we measured the effect of the GE interaction on the association among agronomic traits related to grain yield in soybeans, in relation to the direct and indirect effects obtained in a path analysis; we also evaluated its consequences in the selection for multiple traits.

MATERIAL AND METHODS

The experiments were carried out in three locations in the state of Minas Gerais (Lavras, Ijaci and Itutinga), Brazil, during the 2014/15 and 2015/16 crop seasons. Thirty-five commercial soybean cultivars (Table 1) were investigated; they were chosen because of their potential for production in the southern region of Minas Gerais, Brazil. Experiments were conducted in a randomized complete block design with three replications. Each plot was constituted of two rows 5 m long with 0.5 m spacing between rows.

Table 1. Soybean cultivars evaluated, including maturity group (MG) and growth habit (GH) information.

Cultivars	MG	GH	Cultivars	MG	GH
FPS Iguaçú	5.0	I	5D 690 RR	6.9	I
NS 5106 IPRO	5.1	I	NS 6909 IPRO	6.9	I
NS 5151 IPRO	5.1	I	NS 7000 IPRO	7.0	I
95R51	5.5	I	CD 238 RR	7.1	D
CD 250 RR	5.5	I	M 7110 IPRO	7.1	I
FPS Paranapanema RR	5.6	SD	TMG 716 RR	7.1	I
FPS Solimões RR	5.7	I	97R21	7.2	I
FPS Atlanta	5.8	I	NS 7209 IPRO	7.2	I
RK 5813 RR	5.8	I	CG 68 RR	7.3	I
FPS Júpiter RR	5.9	I	NS 7300 IPRO	7.3	I
NS 5909 IPRO	5.9	I	NS 7338 IPRO	7.3	I
NS 5959 IPRO	5.9	I	CG 67 RR	7.4	SD
5D 615 RR	6.1	I	CG 7464 RR	7.4	SD
FPS Urano RR	6.2	D	CG 7665 RR	7.6	SD
FPS Netuno RR	6.3	I	5G 770 RR	7.7	I
FPS Solar IPRO	6.3	I	CG 8166 RR	7.7	I
FPS Antares RR	6.8	I	5G 830 RR	8.3	D
RK 6813 RR	6.8	I	-	-	-

MG: Maturity Group (Alliprandini et al., 2009); GH: Growth Habit; I: Indeterminate; SD: Semi-determinate; D: Determinate.

The following traits were evaluated: grain yield (Yield): determined from the harvest of each plot, corrected for 13% moisture, (bags.ha⁻¹ - bags of 60 kg); lodging score (LOD.); the lodging score was estimated according to the scale proposed by Bernard et al. (1965); weight of 100 grains (W100) expressed in grams (g); full maturity (AM): determined based on the number of days counted from the seedling emergence date to the date when 95% of the pods are mature (R8 stage); number of pods per plant (NP), number of grains per plant (NG) and number of grains per pods (NGP): five plants were collected at random and the pods and grains were counted manually, obtaining the average number of grains per plant, pods per plant and grains per pod; plant height (HEI) and height of first pod insertion (INP): expressed in centimeters (cm) and obtained by the average of five plants taken at random.

The data were submitted to variance analysis following the strategies presented in Table 2, which were performed in R[®] environment (R Core Team, 2015) and the averages were

grouped by the Scott and Knott (1974) test at 5% probability. Experimental accuracy was measured by estimating the coefficient of variation (CV)(Pimentel-Gomes, 2009) and selective accuracy (Resende and Duarte, 2007).

Table 2. Data analysis strategies for the evaluation of soybean cultivars in Lavras/MG, Ijaci/MG, Itutinga/MG in the crop seasons 2014/15 and 2015/16.

Analysis	Statistical model
Individual –in each site and each crop year	$Y_{ij} = \mu + t_i + b_j + e_{ij}$
Joint of sites in each crop year	$Y_{ijl} = \mu + t_i + b_{j(l)} + a_l + (ta)_{il} + e_{ij(l)}$
Joint of crop years in each site	$Y_{ijk} = \mu + t_i + b_{j(k)} + s_k + (ts)_{ik} + e_{ijk}$
Joint – Total	$Y_{ijkl} = \mu + t_i + b_{j(kl)} + s_k + a_l + (ts)_{ik} + (ta)_{il} + (sa)_{kl} + (tsa)_{ikl} + e_{ijkl(l)}$

Where: $Y_{ij} = Y_{ijl} = Y_{ijk} = Y_{ijkl}$: phenotypic value; μ : constant associated with all observations; t_i : effect of the cultivar i ; b_j : effect of the block j ; $b_{j(l)}$: effect of the block j within the site l ; $b_{j(k)}$: effect of the block j within the crop year k ; $b_{j(kl)}$: effect of the block j within the site l , within the crop year k ; a_l : effect of the site l ; s_k : effect of the crop year k ; $(ta)_{il}$: effect of cultivar-site interaction; $(ts)_{ik}$: effect of cultivar-crop year interaction; $(sa)_{kl}$: effect of site-crop year interaction; $(tsa)_{ikl}$: effect of cultivar-crop year-site interaction; $e_{ij} = e_{ij(l)} = e_{ijk} = e_{ijkl(l)}$: error associated to $Y_{ij} = Y_{ijl} = Y_{ijk} = Y_{ijkl}$, being $e_{ij} = e_{ij(l)} = e_{ijk} = e_{ijkl(l)} \sim N(0, \delta_e^2)$.

The total sum of squares was partitioned into its components in order to estimate the magnitude of the effects of genotype, environment, and genotype x environment interactions on all evaluated traits, according to the estimator:

$$\%FV = \frac{SQ(FV)}{SQT - (SQB + SQR)} \times 100 \quad (\text{Eq. 1})$$

where:

$\%FV$: the magnitude in percentage of effect of the desired source of variation;

$SQ(FV)$: the sum of squares of the desired variation source;

SQT : the total sum of squares;

SQB : the sum squares (S.S) of blocks;

SQR : the sum of squares of the residue.

Path analysis was performed with GENES software (Cruz, 2013), following the same strategies used in the variance analysis. The correlation matrices among the traits were estimated and their significance evaluated by the Mantel test.

Multicollinearity was tested based on the matrix condition number (Montgomery and Peck, 1981). This method considers the condition number (CN) obtained by the ratio of the largest eigen value by the least eigen value of the correlation matrix. If CN is less than 100, multicollinearity is considered weak. In the case of moderate ($100 < CN < 1000$) or strong multicollinearity ($CN > 1000$), it is suggested the elimination of variables from the regression model, preferably those that contribute most to multicollinearity, or the addition of a constant k on the diagonal of the matrix $X'X$, as well as in the ridge regression method (Lima et al., 2014).

The analysis was carried out with a causal diagram of one chain aiming to unfold the correlations in direct and indirect effects of explanatory variables on the grain yield main variable. The decomposition of the correlation between the explanatory variables and the basic variable is given by Cruz (2006):

$$r_{iy} = p_i + \sum_{j \neq i}^n p_j r_{ij} \quad (\text{Eq. 2})$$

where:

r_{iy} : the correlation between the main variable (y) and the i-th explanatory variable;

p_i : the measure of the direct effect of variable i on the main variable;

$p_j r_{ij}$: the measure of the indirect effect of variable i via variable j on the main variable.

Estimates of the model's coefficient of determination and the effect of the residual variable on the main variable were also obtained in this analysis.

The sum of ranks index (SRI) of the evaluated cultivars was used based on its average performance in all the evaluated places and years (Mulamba and Mock, 1978). The characteristics of grain yield, insertion of first pod (INP), plant height (HEI), lodging score (LOD.), weight of 100 grains (W100), number of pods per plant (NP), number of grains per pod (NGP) and full maturity. The SRI was calculated in three ways. The first one uses the same weight for all the evaluated traits, the second uses the correlation value among these traits and grain yield (Resende, 2007) as economic weight of characters, and the third uses the value of the direct effect of traits on yield as economic weight, according to the expression:

$$ISP_j = \sum_{i=1}^n u_i P_{ij} \quad (\text{Eq. 3})$$

where:

ISP_j : the sum of ranks associated with cultivar j;

P_{ij} is the rank obtained for the cultivar j considering the character i;

u_i is the economic weight attributed to character i, being:

$$u_i = \frac{r_{i,y}}{\sum_{i=1}^n r_{i,y}} \text{ or } u_i = \frac{d_{i,y}}{\sum_{i=1}^n d_{i,y}} \quad (\text{Eq. 4})$$

where:

$r_{i,y}$: the correlation between the character i and the basic trait grain yield;

$d_{i,y}$: the direct effect of i on the basic trait grain yield.

With the aid of the statistical software GENES, the coincidence index proposed by Hamblin and Zimmermann (1986) was calculated, with selection intensity of 5, 10, 15, 20, 25, and 30%, in order to test the coincidence of superior cultivars comparing average yield in bags/ha, the sum of ranks index considering the correlation value among the evaluated traits as weight and the grain yield and the sum of ranks index considering the value of the direct effect of evaluated characters on yield as weight, besides the sum of ranks index considering the same weight for all evaluated characters, as follows:

$$IC = \frac{A-C}{M-C} \times 100 \quad (\text{Eq. 5})$$

where:

A: the number of coincident cultivars in the different indexes;

C: the number of coincident cultivars in the different indexes at random ($C = i\% \times M$);

M: the number of superior cultivars selected according to selection index.

RESULTS

We adopted two different strategies to check for accuracy; these being the coefficient of variation (CV) (Pimentel-Gomes, 2009) and the selective accuracy (Resende and Duarte, 2007). In the individual analyses, by site and crop year, good experimental accuracy and high selective accuracy were observed for most of the traits. The selective accuracy ranged from 47.58% (weight of 100 grains in Itutinga 15/16) to 97% (plant height in Lavras 15/16). In general, the greatest magnitude was observed in the estimation of CV for the lodging score.

Significant differences were observed among cultivars for most of the traits, both in the individual analyses by environment and in the joint analyses by site, crop season and total.

The effect of the environment on the expression of most of the traits in the different analysis strategies was also examined. Based on the effect of GE interaction results, we found that the cultivars did not show coincident behavior in the different environments.

The decomposition of the sum of squares into its components, considering all the evaluated environments, allowed estimating the magnitude of the effects of GE, environment and genotype interaction on all evaluated traits. For most of them, a great magnitude of the effect of site and crop years was observed, demonstrating the influence of non-genetic factors on phenotypic expression. For grain yield, e.g., a large magnitude was observed in the effect of site. On the other hand, more than 40% of the phenotypic variation observed in the weight of 100 grains can be attributed to the effect of genotypes. It is also evident that the greater contribution of the cultivar and crop year (C x CYr) interaction was detected for most of the traits (Table 3). As observed in Table 3, the most influential character of the GE interaction was the number of grains per pod (48.61%). Plant height was the character with the lowest effect of GE interaction (4.93%). Considering grain yield, it was observed that 31.7% of the variation resulted from the effects of the GE interaction, which was superior to the effect of genotype (24.29%).

Focusing on the evaluation of multiple traits, statistical methods allow a better understanding of the interrelationship of characters. However, multicollinearity problems may occur and hence impair inferences. When performing the test suggested by Montgomery and Peck (1981), strong multicollinearity (CN > 1000) was found. In this case, one way to minimize the problem is to remove redundant characters.

When analyzing the correlations among characters, there was a high correlation between the number of grains per plant and the number of pods (0.96). Thus, the number of grains per plant was removed. The subsequent test revealed weak multicollinearity.

Table 3. Percentage of the effects of cultivar, site, crop year, and the interactions among these factors on the evaluated characters, relative to the evaluation of soybean cultivars in Lavras/MG, Ijaci/MG and Itutinga/MG in 2014/15 and 2015/16 crop years. Lavras – MG, 2017.

Traits	Cultivars	Sites	Crop years	SxCYr	CxS	CxCYr	CxSxCYr
Yield ¹	24.29	23.76	5.05	13.88	9.36	14.21	8.13
NP ²	35.64	25.11	0.07	1.83	13.47	12.20	10.14
NG ³	31.84	17.62	1.77	2.32	14.12	16.53	12.81
NGP ⁴	17.99	9.27	10.00	5.55	16.35	17.73	14.53
W100 ⁵	42.61	8.61	14.83	0.51	10.19	9.51	10.83
HEI ⁶	29.36	12.13	40.42	12.17	1.17	1.92	1.84
INP ⁷	35.29	22.96	12.55	14.67	5.42	3.53	4.75
LOD ⁸	31.76	1.96	19.77	5.10	7.67	17.72	9.81
AM ⁹	33.46	26.33	13.15	9.17	7.77	5.03	3.80

Yield: grain yield, bags/ha; NP: number of pods per plant; NG: number of grains per plant; NGP: number of grains per pod; W100: weight of 100 grains, g; HEI: plant height, cm; INP: height of first pod insertion, cm; LOD: plant lodging; AM: full maturity, days; SxCYr.: sites x crop years interaction; CxS: cultivars x sites interaction; CxCYr: cultivars x crop years interaction; CxSxCYr.: cultivars x sites x crop years interaction.

When analyzing the correlations between grain yield and the other traits, we noted a significant change in the magnitude and direction among the different analysis strategies, i.e., individual and joint analysis by crop season, site and total. The highest values of positive and significant correlation with grain yield were observed for the traits height of first pod insertion (0.6813) and full maturity (0.6990) (Table 4).

Table 4. Correlation among characters and grain yield considering different analysis strategies relative to the evaluation of soybean cultivars in Lavras/MG, Ijaci/MG and Itutinga/MG in 2014/15 and 2015/16 crop years. Lavras – MG, 2017.

Variable	Correlations between Yield											
	Joint	Site 1	Site 2	Site 3	CYr.1	CYr.2	CYr.1S1	CYr.1S2	CYr.1S3	CYr.2S1	CYr.2S2	CYr.2S3
NP ¹³	0.3674*	-0.0985	0.4607*	0.4075*	0.4984*	-0.1929	-0.2384	0.6672*	0.5051*	-0.1950	-0.0534	-0.1196
NG ¹⁴	0.4190*	-0.0097	0.4436*	0.4518*	0.5661*	-0.0868	-0.2134	0.6875*	0.6209*	0.0661	-0.0080	-0.0047
NGP ¹⁵	-0.0165	0.1452	-0.1062	-0.1316	0.3737*	0.2015	0.0490	0.2639	0.3992*	0.3417*	0.0887	0.0637
W100 ¹⁶	0.1220	0.1966	0.2772*	-0.0015	0.0503	0.2857*	0.1925	0.1637	-0.0459	0.3426*	0.3228*	0.1947
HEI ¹⁷	0.4914*	0.2004	0.5763*	0.4519*	0.7266*	-0.0521	0.3817*	0.7227*	0.7425*	-0.0287	0.2421	-0.1935
INP ¹⁸	0.6813*	0.2401	0.7005*	0.6686*	0.7986*	0.2522*	0.4395*	0.6596*	0.7706*	-0.0473	0.4577*	0.3319*
LOD ¹⁹	0.3505*	0.0145	0.4771*	0.2342	0.4099*	-0.2413	0.0243	0.3236	0.3442*	-0.0675	-0.0472	-0.3083
AM ²⁰	0.6990*	0.1404	0.8070*	0.6200*	0.7884*	0.2124	0.0348	0.8227*	0.6918*	0.0975	0.3946*	0.1073

*Significant at the 5% level by the Mantel test. Joint: Itutinga, Lavras and Ijaci in 2014/15 and 2015/16 crop years; Site 1: Itutinga in 2014/15 and 2015/16 crop years; Site 2: Lavras in 2014/15 and 2015/16 crop years; Site 3: Ijaci in 2014/15 and 2015/16 crop years; CYr.1: Itutinga, Lavras and Ijaci in 2014/15 crop year; CYr.2: Itutinga, Lavras and Ijaci in 2015/16 crop year; CYr.1S1: Itutinga in 2014/15 crop year; CYr.1S2: Lavras in 2014/15 crop year; CYr.1S3: Ijaci in 2014/15 crop year; CYr.2S1: Itutinga in 2015/16 crop year; CYr.2S2: Lavras in 2015/16 crop year; CYr.2S3: Ijaci in 2015/16 crop year; NP: number of pods per plant; NG: number of grains per plant; NGP: number of grains per pod; W100: weight of 100 grains, g; HEI: plant height, cm; INP: height of first pod insertion, cm; LOD: plant lodging; AM: full maturity in days.

The highest values of direct effects on grain yield were observed for full maturity and insertion of the first pod, considering all the evaluated sites and crop season (Table 5).

Table 5. Direct effect of the explanatory variables on grain yield considering the different analysis strategies relative to the evaluation of soybean cultivars in Lavras/MG, Ijaci/MG and Itutinga/MG in 2014/15 and 2015/16 crop years. Lavras – MG, 2017.

Variable	Direct effect on grain yield											
	Joint	Site 1	Site 2	Site 3	CYr.1	CYr.2	CYr.1S1	CYr.1S2	CYr.1S3	CYr.2S1	CYr.2S2	CYr.2S3
NP ¹³	-0.0520	-0.2068	0.2254	0.1762	-0.3453	-0.0812	-0.2812	0.0735	0.1386	-0.2143	0.1717	0.0576
NGP ¹⁴	0.3002	0.1360	0.0608	0.1536	0.2460	0.3975	0.0844	0.0586	0.1279	0.3235	0.3426	0.1460
W100 ¹⁵	0.1383	0.2366	0.3536	0.0957	0.1759	0.0264	0.2734	0.3344	0.2889	0.2284	0.2092	0.0709
HEI ¹⁶	0.0983	0.1632	0.0202	0.0907	0.5082	-0.2187	0.4588	0.3397	0.4147	0.0252	0.2003	-0.2103
INP ¹⁷	0.3798	0.1376	0.3180	0.4472	0.1592	0.4851	0.2622	0.0120	0.1986	-0.2934	0.5392	0.4420
LOD ¹⁸	-0.4523	-0.0822	-0.2629	-0.3364	-0.2238	-0.5770	-0.1360	0.1305	0.0473	0.1604	-0.8629	-0.2976
AM ¹⁹	0.8509	0.2208	0.6302	0.4087	0.6746	0.6469	-0.1179	0.4991	0.2644	0.4985	0.6158	0.1238
R ²	0.7361	0.1822	0.8130	0.5660	0.8652	0.5137	0.4067	0.8255	0.7679	0.2815	0.6677	0.3086
Residual Variable	0.5137	0.9043	0.4325	0.6588	0.3671	0.6974	0.7702	0.4177	0.4817	0.8477	0.5765	0.8315

Joint: Itutinga, Lavras and Ijaci in 2014/15 and 2015/16 crop years; Site 1: Itutinga in 2014/15 and 2015/16 crop years; Site 2: Lavras in 2014/15 and 2015/16 crop years; Site 3: Ijaci in 2014/15 and 2015/16 crop years; CYr.1: Itutinga, Lavras and Ijaci in 2014/15 crop year; CYr.2: Itutinga, Lavras and Ijaci in 2015/16 crop year; CYr.1S1: Itutinga in 2014/15 crop year; CYr.1S2: Lavras in 2014/15 crop year; CYr.1S3: Ijaci in 2014/15 crop year; CYr.2S1: Itutinga in 2015/16 crop year; CYr.2S2: Lavras in 2015/16 crop year; CYr.2S3: Ijaci in 2015/16 crop year; NP: number of pods per plant; NGP: number of grains per pod; W100: weight of 100 grains, g; HEI: plant height, cm; INP: height of first pod insertion, cm; LOD: plant lodging; AM: full maturity in days.

However, for most of the analysis strategies, the direct effects of the evaluated traits were inferior to the residual effect, demonstrating that indirect selection through these traits may not result in gains. Furthermore, the low coefficients of determination obtained in these analysis strategies indicate the need to include new traits in the causal diagram.

A useful methodology for identifying and selecting superior cultivars that associates good attributes with several characters, is the use of index selection. In our study, we adopted the sum of ranks index (SRI). When comparing the ranking of cultivars by yield as

well as by this index, considering the different economic weights (Table 6), we observed that there is great similarity between grain yield classification and the SRI that uses the direct effect as economic weight.

Table 6. Ranking of the cultivars by grain yield and by sum of ranks index relative to the evaluation of soybean cultivars in Lavras/MG, Ijaci/MG and Itutinga/MG in 2014/15 and 2015/16 crop years. Lavras – MG, 2017.

Cultivars	Yield (bags/ha)	SRI	SRI (correlation)	SRI (direct effect)
NS7000 IPRO	1	5	2	2
NS 7209 IPRO	2	16	20	1
NS 7300 IPRO	3	27	21	3
CG 68 RR	4	11	7	5
CG 7464 RR	5	10	5	4
CD 238 RR	6	20	13	7
FPS Antares RR	7	8	1	6
NS 7338 IPRO	8	24	27	9
CG 8166 RR	9	34	30	11
FPS Netuno RR	10	12	4	8
CG 7665 RR	11	31	29	13
RK 6813 RR	12	17	11	12
97R21	13	28	22	17
RK 5813 RR	14	13	3	15
M7110 IPRO	15	2	8	10
FPS Solar RR	16	9	14	16
CG 67 RR	17	21	24	19
NS 6909 IPRO	18	6	16	14
5G 830 RR	19	29	32	20
NS 5959 IPRO	20	3	10	18
5G 770 RR	21	33	28	23
FPS Atlanta IPRO	22	7	15	21
NS 5909 IPRO	23	1	6	22
5D 690 RR	24	32	35	25
NS 5151 IPRO	25	4	12	24
5D 615 RR	26	19	17	26
FPS Júpiter RR	27	22	25	28
FPS Paranapanema RR	28	18	9	29
FPS Urano RR	29	26	31	27
TMG 716 RR	30	35	33	35
FPS Solimões RR	31	30	26	33
NS 5106 IPRO	32	15	18	30
95R51	33	23	23	32
CD 250	34	25	34	34
FPS Iguaçu RR	35	14	19	31

Yield (bags/ha): ranking by average grain yield in bags/ha; SRI: ranking by sum of ranks index considering the same weight for all evaluated characters; SRI (correlation): ranking by sum of ranks index considering as weight the correlation value among the evaluated characters and grain yield; SRI (direct effect): ranking by sum of ranks index considering as weight the direct effect value of the evaluated characters on grain yield.

The coincidence index was obtained to verify the similarity in the classification of cultivars, comparing the different used strategies and disregarding chance. In Table 7, it can be observed that from the selection intensity of 15%, there is a greater coincidence between the rankings by yield and by the SRI, which uses the direct effect as economic weight in the selection of superior cultivars. This coincidence reaches 100% with a selection index of 30%, while the maximum coincidence with yield is 42.86% in the other comparisons.

Table 7. Coincidence index between the ranking by grain yield and by sum of ranks index from different selection intensities relative to the evaluation of soybean cultivars in Lavras/MG, Ijaci/MG and Itutinga/MG in 2014/15 and 2015/16 crop years. Lavras – MG, 2017.

Selection Intensity (%)	Coincidence Index					
	Yield x SRI	Yield x SRI (correlation)	Yield x SRI (direct effect)	SRI x SRI (correlation)	SRI x SRI (direct effect)	SRI (correlation) x SRI (direct effect)
5	0.00	0.00	0.00	0.00	100.00	0.00
10	0.00	33.33	33.33	33.33	33.33	66.67
15	0.00	20.00	80.00	40.00	20.00	40.00
20	28.57	42.86	85.71	71.43	28.57	28.57
25	25.00	37.50	87.50	75.00	25.00	25.00
30	30.00	40.00	100.00	80.00	30.00	40.00

Yield: ranking by average grain yield in bags/ha; SRI: ranking by sum of ranks index considering same weight to all evaluated characters; SRI (correlation): ranking by sum of ranks index considering as weight the correlation value among the evaluated characters and grain yield; SRI (direct effect): ranking by sum of ranks index considering as weight the direct effect value of the evaluated characters over grain yield.

DISCUSSION

To achieve success in agricultural experimentation, high experimental accuracy is desirable, which ensures more accurate estimates and hence more reliable results. The CV value depends only on the residual variation as a ratio of the experiment average (Resende and Duarte, 2007). Thus, higher CV values are expected for attributes with lower averages (Soares et al., 2015). This is a possible explanation for the low accuracy associated with lodging score.

Individual analyses are generally less accurate when compared to the joint analysis. This fact can be justified by the use of a greater number of replications when performing a joint analysis, which provides greater experimental accuracy and lower error, associated to the estimates (Pimentel-Gomes, 2009).

In other studies, a significant difference was also observed among soybean cultivars in our region (Gesteira et al., 2015). This variation in cultivar behavior can be explained mainly by the genetic background of the cultivars, i.e., differences in full maturity and growth habit, besides resistance to different pathogens. The existence of variation among the cultivars opens the possibility of selection and quantification of possible correlations among characters.

The environmental effect is due to the combination of predictable and unpredictable factors (Allard and Bradshaw, 1964), since the cultivars were tested in two crop years and in different locations in the state of Minas Gerais, Brazil. The effect of GE interaction observed in most of the traits in the joint analyses by site, crop season and total results from the combination of environmental factors associated with cultivars. Other studies have been conducted for soybean evaluation in the state of Minas Gerais, Brazil that reported GE interactions (Gesteira et al., 2015; Soares et al., 2015; Silva et al., 2016; Silva et al., 2017).

The results observed for the decomposition of the sum of squares into its components are similar to those of other studies (Gurmu et al., 2009; Salmeron et al., 2014) where the effect of GE interaction overcomes the effect of genotype on grain yield variation.

Considering the differences observed in the magnitude of correlations between grain yield and the other traits, a probable explanation for the change in the correlations

among traits in soybean is the GE interaction. In this case, the environmental factor can be attributed to the differences of climate and soil between the sites and crop years.

Nogueira et al. (2012) and Silva et al. (2016) also found a positive and significant correlation between grain yield and height of first pod insertion in soybean (0.23 and 0.579, respectively). Leite et al. (2015) also report a positive correlation between these two traits (0.252), although, not significant. However, Teodoro et al. (2015) identified a negative and non-significant phenotypic correlation between these traits, with a value of -0.3275 for cultivars with indeterminate growth and -0.2758 for cultivars with determinate growth. In relation to the correlation between grain yield and full maturity, Peluzio et al. (2005) found a positive and non-significant estimation among these traits (0.6091). Gesteira et al. (2015) reported a positive and high association between grain yield and full maturity, as well as a negative and high association between yield and protein content, full maturity and protein content, and between oil and protein contents in grains.

For breeding purposes, it is important to identify among the high correlations with the main character those with the greatest direct effect on selection, so that the correlated response through indirect selection is efficient (Cruz and Carneiro, 2012). These estimates can be obtained through path analysis.

Studies performed by Iqbal et al. (2003) and Nogueira et al. (2012) showed a higher direct effect on grain yield for the number of pods per plant, weight of 100 grains and number of grains per pod. Teodoro et al. (2015) found that in determinate growth cultivars, the number of pods per plant and number of branches had a greater direct effect on grain yield. For cultivars with indeterminate growth, besides these two traits, plant height also has a high direct effect on grain yield.

The difference observed in the magnitude of direct effects of traits on grain yield, when considering the different analysis strategies, shows the influence of sites and crop season on the association among traits. The importance of performing the evaluation of genotypes in different sites and crop season is therefore emphasized in order to obtain more accurate results, especially when the indirect selection of traits is desired.

We found that index selection was efficient in the classification of the best cultivars. Index selection contemplates more traits and yield is a useful trait to identify the best genotypes. Furthermore, the use of the direct effect may be more efficient in the selection of superior cultivars when compared to the use of correlation among traits, as suggested by Resende (2007). As there were differences in the classification of cultivars among the different strategies, we suggest that the use of only one statistical method may not be as efficient as joint analysis strategies to identify superior cultivars.

ACKNOWLEDGMENTS

We thank the Federal University of Lavras (UFLA), the Graduate Program in Genetics and Plant Breeding, the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Fundação de Amparo à Pesquisa do Estado de Minas Gerais

(FAPEMIG) and the Comissão de Aperfeiçoamento de Pessoal do Nível Superior (CAPES) for their support.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Allard RW and Bradshaw AD (1964). Implications of genotype-environmental interactions in applied plant breeding. *Crop Sci.* 4: 503-508.
- Alliprandini L, Abatti C, Bertagnolli PF, Cavassim JE, et al. (2009). Understanding soybean maturity groups in Brazil: environment, cultivar classification, and stability. *Crop Sci.* 49: 801-808.
- Bernard RL, Chamberlain DW and Lawrence RE (1965). Results of the cooperative uniform soybean tests. USDA, Washington.
- Baraskar VV, Vachhani JH, Kachhadia VH, Patel MB, et al. (2015). Correlation and path analysis for seed yield in soybean [*Glycine max* (L.) Merrill]. *Electron. J. Plant Breed.* 6: 570-573.
- Bizari EH, Val BHP, Pereira EM, Di Mauro AO, et al. (2017). Selection indices for agronomic traits in segregating populations of soybean. *Cienc. Agron.* 48: 110-117.
- Cruz CD (2006). Programa GENES: Estatística experimental e matrizes. Editora UFV. Viçosa.
- Cruz CD (2013). GENES - A software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35: 271-276.
- Cruz CD and Carneiro PCS (2014). Modelos Biométricos aplicados ao melhoramento genético. vol. 2.3ª edição. Editora UFV. Viçosa.
- Cruz CD, Regazzi AJ and Carneiro PCS (2012). Modelos biométricos aplicados ao melhoramento genético. vol. 1. Editora UFV. Viçosa.
- Gesteira GS, Zambiazzi EV, Bruzi AT, Soares IO, et al. (2015). Seleção fenotípica de cultivares de soja precoce para a região Sul de Minas Gerais. *Agro. geoamb.* 7: 79-88.
- Gurmu F, Mohammed H and Alemaw G (2009). Genotype x environment interactions and stability of soybean for grain yield and nutrition quality. *Afr. Crop Sci. J.* 17: 87-99.
- Hamblin J and Zimmermann MJO (1986). Breeding common bean for yield mixtures. *Plant Breeding Rev.* 4: 245-272.
- Iqbal S, Mahmood T, Ali M, Anwar M, et al. (2003). Path coefficient analysis in different genotypes of soybean (*Glycine max* (L.) Merrill). *Pak. J. Biol. Sci.* 12: 1085-1087.
- Leite WS, Pavan BE, Matos Filho CHA, Feitosa FS, et al. (2015). Estimativa de parâmetros genéticos e correlações entre caracteres agrônômicos em genótipos de soja. *Nativa.* 3: 241-245.
- Lima RP, Silva AR and Oliveira DMS (2014). Análise de trilha de atributos físicos na resistência à penetração de umlato solo amarelo. *Rev. Agric. Neotrop.* 1: 65-74.
- Lopes ACA, Vello NA, Pandini F, Rocha MM, et al. (2002). Variabilidade e correlações entre caracteres em cruzamentos de soja. *Scientia Agricola.* 59: 341-348.
- Montgomery DC and Peck EA (1981). Introduction to linear regression analysis. John Wiley & Sons, New York.
- Mulamba NN and Mock JJ (1978). Improvement of yield potential of the method Eto Blanco maize (*Zea mays* L.) population by breeding for plant traits. *Egypt J. Genet. Cytol.* 7: 40-51.
- Neto AF, Gravian GA, Monteiro MMS, Morais FB, et al. (2011). Análise de trilha do rendimento de grãos de soja na microrregião do Alto Médio Gurguéia. *Com. Sci.* 2: 107-112.
- Nogueira APO, Sedyama T, Sousa LB, Hamawaki OT, et al. (2012). Análise de trilha e correlações entre caracteres em soja cultivada em duas épocas de semeadura. *Biosci. J.* 28: 877-888.
- Peluzio JM, Almedia RD, Fidelis RR, Almedia Júnior D, et al. (2005). Correlação entre caracteres de soja, em Gurupi, Tocantins. *Ceres.* 52: 779-786.
- Perini LJ, Júnior NSF, Destro D and Prete CEC (2012). Componentes da produção em cultivares de soja com crescimento determinado e indeterminado. *Semin. Cienc. Agrar.* 33: 2531-2544.
- Pimentel-Gomes F (2009). Curso de estatística experimental. 15ª edição. FEALQ, Piracicaba.
- R Core Team (2015). R: A language and environment for statistical computing, version 3.2.2. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org> (accessed January 15, 2017).
- Reis CAF, Gonçalves FMA, Ramalho MAP and Rosado AM (2015). Estratégias na seleção simultânea de vários caracteres no melhoramento do *Eucalyptus*. *CiFl.* 25: 457-467.
- Resende MDV and Duarte JB (2007). Precisão e controle de qualidade em experimentos de avaliação de cultivares. *Pesqui. Agropecu. Trop.* 3: 182-194.

- Resende MDV (2007). Matemática e estatística na análise de experimentos e no melhoramento genético. Embrapa Florestas, Colombo.
- Salmeron M, Gbur EE, Bourland FM, Buehring NW, et al. (2014). Soybean maturity group choices for early and late plantings in the Midsouth. *Agron. J.* 5: 1893-1901.
- Santos FS, Amaral Júnior AT, Freitas Júnior SD, Rangel RM, et al. (2007). Predição de ganhos genéticos por índices de seleção na população de milho-pipoca UNB-2U sob seleção recorrente. *Bragantia.* 66: 389-396.
- Scott AJ and Knott M (1974). A Cluster analysis method for grouping means in the analysis of variance. *Biometrics.* 30: 507-512.
- Silva KB, Bruzi AT, Zuffo AM, Zambiazzi EV, et al. (2016). Adaptability and phenotypic stability of soybean cultivars for grain yield and oil content. *Genet. Mol. Res.* 15: 2.
- SilvaKB, Bruzi AT, Zambiazzi EV, Soares IO, et al. (2017). Adaptability and stability of soybean cultivars for grain yield and seed quality. *Genet. Mol. Res.* 15: 1-15.
- Soares IO, Rezende PM, Bruzi AT, Zambiazzi EV, et al. (2015). Adaptability of soybean cultivars in different crop years. *Genet. Mol. Res.* 14: 8995-9003.
- Teodoro PE, Ribeiro LP, Corrêa CCG, Luz Júnior RAA, et al. (2015). Path analysis in soybean genotypes as function of growth habit. *Biosci. J.* 3: 794-799.