

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/326044534>

# SELECTION OF EARLY SOYBEAN INBRED LINES USING MULTIPLE INDEX

**Article** in *Crop Science* · June 2018

DOI: 10.2135/cropsci2018.05.0295

CITATIONS

0

READS

68

**5 authors**, including:



**Gabriel Gesteira**

Escola Superior de Agricultura "Luiz de Queiroz"

**4** PUBLICATIONS **10** CITATIONS

[SEE PROFILE](#)



**Adriano Bruzi**

Universidade Federal de Lavras (UFLA)

**56** PUBLICATIONS **142** CITATIONS

[SEE PROFILE](#)



**Roberto Kazuhiko Zito**

Brazilian Agricultural Research Corporation (EMBRAPA)

**5** PUBLICATIONS **15** CITATIONS

[SEE PROFILE](#)



**Vanoli Fronza**

Brazilian Agricultural Research Corporation (EMBRAPA)

**19** PUBLICATIONS **113** CITATIONS

[SEE PROFILE](#)

**Some of the authors of this publication are also working on these related projects:**



SELECTION FOR MULTIPLE TRAITS IN UPLAND RICE PROGENIES USING THE Z INDEX [View project](#)

## SELECTION OF EARLY SOYBEAN INBRED LINES USING MULTIPLE INDEX

**Gabriel de Siqueira Gesteira<sup>\*</sup>, Adriano Teodoro Bruzi, Roberto Kazuhiko Zito, Vanoli Fronza and Neylson Eustáquio Arantes**

Gabriel de Siqueira Gesteira, Genetics and Plant Breeding Department, Luiz de Queiroz College of Agriculture (ESALQ/USP), Piracicaba, SP, Brazil; Adriano Teodoro Bruzi, Plant Sciences Department, Federal University of Lavras (UFLA), Lavras, MG, Brazil; Roberto Kazuhiko Zito, EMBRAPA Soybean, Londrina, PR, Brazil; Vanoli Fronza, EMBRAPA Wheat, Passo Fundo, RS, Brazil; Neylson Eustáquio Arantes, Triangle Foundation, Uberaba, MG, Brazil. Received \_\_\_\_\_. <sup>\*</sup> Corresponding author ([gabrielgesteira@usp.br](mailto:gabrielgesteira@usp.br)).

### ABSTRACT

The soybean crop (*Glycine max* L.) has been studied and enhanced for most of its economically important traits. Previous research has studied the association among them and the effect of the genotype by environment interaction (GEI); however, less is known about their correlation considering the absolute maturity, as well as the use of multiple selection indexes to study the GEI and select superior cultivars. Regarding this, the aim of the present study was to identify lines that associate precocity, good yield performance and high oil and protein contents in the grains, as well as to estimate the correlation among these traits and study the effect of GEI, using a standardized multiple selection index. Trials were conducted in two crop seasons in the state of Minas Gerais, Brazil, with 39 lines in 13 evaluation environments. The experiments were conducted in a randomized complete block design with 4 replications, and the grain yield, absolute maturity, and oil and protein contents in the grains were evaluated. The results indicated high experimental precision and accuracy, with significant differences among lines for all traits. High magnitude correlations between evaluated traits were found, highlighting the negative correlation between absolute maturity and protein content in the grains. The GEI was also significant, and the use of the multiple selection index was efficient to identify superior and stable inbred lines by the GGE Biplot method, which explained 82.23% of the GEI effect. Lines 27 and 31 stood out from the others because they associated stability and good performance for all evaluated traits.

**Keywords:** Correlation; GEI; GGE Biplot; *Glycine max*; Grain yield.

## INTRODUCTION

The soybean (*Glycine max* L.) is one of the most important crops in the international market, being the fourth most produced and consumed crop worldwide. It has average oil and protein contents of 20 and 40%, respectively, which makes it an excellent raw material for feed preparation, oil extraction, biodiesel production, as well as containing cosmetic, nutraceutical and pharmaceutical properties (EMBRAPA, 2013).

Brazil is the second largest soybean producer worldwide and is one of the largest exporters of its products and by-products (CONAB, 2016). The country stands out in the international market because it shows great potential for agricultural production and for the possibility of harvesting two crops successively in the rainy season (summer) in some regions, which allows for greater production in the same area, and hence greater economic returns.

After the beginning of the second harvest in Brazil, a good part of the soybean breeding programs turned to the development of earlier cultivars, aiming to anticipate the harvest and to enable the planting of the crop in succession. However, this reduction in the crop cycle may influence the expression of other traits, especially those of greater economic interest, such as grain yield, size and architecture of plants, resistance to biotic and abiotic factors, as well as the oil and protein contents in the grains. It is known that there is a positive and large magnitude correlation between grain yield and absolute maturity in soybean crop, which means that the reduction in the crop cycle would imply generally a reduction in the average yield of cultivars.

The estimation of phenotypic, genetic and environmental parameters allows identifying and studying the traits of interest for breeding programs. It is also possible to estimate the correlation among them, aiming to verify if the observed variation in one trait can influence the variation observed in another trait (Cruz et al., 2012). Several studies have been conducted with the aim of estimating these parameters for several traits in the soybean crop, as well as their correlations

(Carvalho et al., 2002; Lopes et al., 2002; Carpentieri-Pipolo et al., 2005; Peluzio et al., 2005; Almeida et al., 2010; Nogueira et al., 2012; Rigon et al., 2012; Gesteira et al., 2015).

Although most of these studies deal with quantitative traits, no studies were found in the literature involving the correlation between oil and protein contents with absolute maturity for the soybean crop. The study of the correlation among these traits is extremely important, since it allows identifying if the change in the absolute maturity of commercial cultivars can affect the oil and protein contents present in the grains.

In value for cultivation and use (VCU) experiments, several traits are usually evaluated. An alternative for contemplating more than one characteristic in the selection is the adoption of an index. This strategy has been applied in the soybean crop in some occasions (Soares et al., 2015; Silva, 2016). However, it should be noted that there are no reports on the use of this tool in the selection of soybean cultivars for absolute maturity, yield, and grain quality.

The genotype by environment interaction is often reported for different traits in the soybean crop (Carvalho et al., 2002; Lima et al., 2008; Vasconcelos et al., 2010; Marques et al., 2011; Barros et al., 2012; Bueno et al., 2013; Silva et al., 2015; Soares et al., 2015) and other crop species, like barley, triticale and wheat (Kendal, 2015; Kendal and Sener, 2015; Dogan et al., 2016; Kendal and Sayar, 2016; Kendal et al., 2016; Kendal and Tekdal, 2016). When the existence of interaction is verified, a strategy that can be adopted is the adaptability and stability analysis through the GGE Biplot method. Several studies report the use of this method, especially for the trait grain yield. However, results on the study of adaptability and stability are not available when considering multiple traits in the soybean crop.

Based on the above, the present study was performed with the aim of: obtaining information on the association among the traits of grain yield, absolute maturity, and oil and protein contents in grains; studying the genotype by environment interaction (crop seasons and sites) in the choice of cultivars using multiple index selection; identifying soybean lines that associate early maturity, good yield and grain quality.

## MATERIAL AND METHODS

### Conduction of the experiments

The experiments were conducted in 10 cities of the Minas Gerais state (Brazil), during two crop seasons, with the aim of covering the soybean production area of the state (Table 1). In the first crop season, 17 elite inbred lines were evaluated, and in the second crop season, 20 elite inbred lines from the same program were evaluated, three of which were common to the first crop season. Five commercial cultivars (Treatments 1 to 5) were used as controls in all experiments, totaling 39 different inbred lines (Table 2). All elite inbred lines evaluated were developed at EMBRAPA Soybean Breeding Program (in partnership with EPAMIG, Triângulo Foundation, EMATER-GO and CTPA).

The no-tillage system was adopted and the seeds were inoculated with *Bradyrhizobium japonicum* using a liquid inoculant, in the proportion of 1,200,000 bacteria per seed. Seeding was performed mechanically in both crop seasons. The thinning was done 25 days after germination, aiming to guarantee the final stand of 15 plants per linear meter. The other cultural practices were performed according to the procedure shown by Soares et al. (2015).

The 39 treatments were evaluated in randomized complete blocks design with four replicates, each plot consisting of four rows of 5 m in length, spaced 0.5 m between rows. The seeds of the two central rows were harvested, except for the initial and terminal 0.5 m of each row, totaling 4 m<sup>2</sup> of useful area.

The following traits were evaluated:

a) grain yield (kg.ha<sup>-1</sup>), obtained by individual harvesting of each plot, weighing and correction for 13% moisture and extrapolation of obtained value for the number of kilograms harvested in one hectare;

b) Absolute maturity, comprising the number of days from sowing until maturity, represented by 95% of plants with mature pods in each plot;

c) The percent protein and oil contents in the grains were determined in whole grains by the technique of the near infrared reflectance spectroscopy (NIRS), according to Heil (2015). The grains of every line were submitted to duplicate readings through Thermo equipment, Antaris II model, with integrating sphere with resolution of  $4\text{ cm}^{-1}$ .

The obtained data were used to perform individual analyzes of variance per site in each crop season, following the model:

$$y_{ij} = \mu + g_i + b_j + \varepsilon_{ij}$$

where:

$y_{ij}$ : phenotypic observation of line  $i$  in block  $j$ ;

$\mu$ : overall average;

$g_i$ : effect of line  $i$ ;

$b_j$ : effect of block  $j$ ;

$\varepsilon_{ij}$ : error associated with the observation of line  $i$  in block  $j$ .

Later, since the data was unbalanced and the analysis was based on a mixed model, a joint deviance analysis was performed considering all the environments (site-year combination), following the model:

$$y_{ijk} = \mu + g_i + a_k + b_{j(k)} + ga_{ik} + \varepsilon_{ijk}$$

where:

$y_{ijk}$ : phenotypic observation of line  $i$  in block  $j$  in the environment  $k$ ;

$\mu$ : overall average;

$g_i$ : effect of line  $i$ , fixed nature;

$a_k$ : effect of environment  $k$ , random nature;

$b_{j(k)}$ : effect of block j in environment k, random nature;

$ga_{ik}$ : effect of the interaction line i and environment k, random nature;

$\varepsilon_{ijk}$ : error associated with the observation of line i in block j in the environment k.

The joint deviance analysis was performed using the restricted maximum likelihood (REML) method, with a significance test of fixed effect by the F value, and the random effects by the likelihood-ratio test (LRT). The averages obtained in all analyses were compared by Scott and Knott (1974) at 5% probability. Experimental accuracy was measured by estimating the coefficient of variation (CV) and accuracy (Resende and Duarte, 2007), according to the estimators:

$$CV = \frac{\sigma}{\bar{y}} \times 100$$

where:

$CV$ : coefficient of variation expressed as a percentage;

$\sigma$ : standard deviation of the trait;

$\bar{y}$ : average of the trait.

$$r = \sqrt{1 - \frac{1}{F_c}} \times 100$$

where:

$r$ : accuracy expressed as a percentage;

$F_c$ : calculated F value.

The coefficient of variation (CV) is dependent on the phenotypic average, and thus the lower the average value, the higher the CV. On the other hand, the accuracy is independent of the average

magnitude, which is the main advantage of using this parameter when evaluating the experimental accuracy (Resende and Duarte, 2007).

### Estimation of phenotypic parameters and index selection

With the adjusted phenotypic averages of the lines obtained in the joint analysis, the association among yield, absolute maturity, and oil and protein contents in the grains was estimated, adopting the following estimator:

$$r_p = \frac{cov_{p_{XY}}}{\sqrt{\sigma_{p_X}^2 \cdot \sigma_{p_Y}^2}}$$

where:

$r_p$ : estimator of the phenotypic correlation between the X and Y traits;

$cov_{p_{XY}}$ : phenotypic covariance between the X and Y traits;

$\sigma_{p_X}^2$ : phenotypic variance for trait X;

$\sigma_{p_Y}^2$ : phenotypic variance for trait Y.

The significance of the obtained correlations were estimated by the t-test with n-2 degrees of freedom, where n corresponds to the number of evaluated genotypes.

The standardized multiple selection index (SMSI) was calculated based primarily on the standardization of the phenotypic averages obtained in the individual analyses for all lines, using the following estimator:

$$Z_{ik} = \frac{\bar{y}_{ik} - \bar{y}_{.k}}{\sigma_{.k}}$$



where:

$Z_{ik}$ : standardized average;

$\bar{y}_{ik}$ : average of line i in the environment k;

$\bar{y}_{.k}$ : overall average in the environment k;

$\sigma_{.k}$ : standard deviation in the environment k.

The index was calculated considering the grain yield, absolute maturity, oil content, and protein content, which were added originating the SMSI, according to the following expression (Mendes et al., 2009):

$$Z_{MIS} = Z_{GY} - Z_{AM} + Z_{Oil} + Z_{Protein}$$

where:

$Z_{MIS}$ : multiple index selection;

$Z_{GY}$ : standardized average for the trait grain yield;

$Z_{AM}$ : standardized average for the trait absolute maturity;

$Z_{Oil}$ : standardized average for the trait oil content;

$Z_{Protein}$ : standardized average for the trait protein content;

The obtained values were added to a constant in order to eliminate negative values.

### **Study of genotype x environment interaction**

The standardized multiple selection index (SMSI) was used to study the effect of genotype x environment interaction, with the objective of including the four main traits associated to line yield

and grain quality (yield, absolute maturity, and oil and protein contents in the grains), thus forming a multivariate analysis.

Due to the high imbalance degree of the data, the phenotypic values of each absent genotype x environment interaction were predicted, aiming to fill in the double entry tables of the SMSI considering all the environments for each evaluated characteristic, as well as to enable the study of genotype x environment interaction. The method used to predict missing values was based on the singular value decomposition (SVD) in combination with the EM algorithm for the iterations in order to find the value with the maximum likelihood of occurrence, as proposed by Yan (2013).

Cross-validation was also performed to calculate the predictive accuracy through resampling and predicting values from a sub-matrix, contemplating the balanced data of the second harvest, after which the correlation of the predicted values with the observed phenotypic values was calculated. The value of the predictive accuracy was calculated based on the average of correlations obtained in 1000 resampling, for all traits.

The double entry tables of the four characteristics of interest, containing the 39 genotypes evaluated in the 12 environments, were filled with the predicted values and used to calculate the SMSI. With the balanced data, the decomposition of genotypic effects and G x E interaction was performed following the model proposed by Yan et al. (2000), according to the equation:

$$Y_{ij} = \mu + \beta_j + \lambda_1 \alpha_{i1} \gamma_{1j} + \lambda_2 \alpha_{i2} \gamma_{2j} + \varepsilon_{ij}$$

where:

$Y_{ij}$ : phenotypic average of line i in the environment j;

$\mu$ : overall average of the experiment;

$\beta_j$ : main effect of the environment j;

$\lambda_1$ : largest eigenvalue of the first principal component (PCA1);

$\alpha_{i1}$ : eigenvector of line i for PCA1;

$\gamma_{1j}$ : eigenvector of the environment  $j$  for PCA1;

$\lambda_2$ : largest eigenvalue of the second principal component (PCA2);

$\alpha_{i2}$ : eigenvector of line  $i$  for PCA2;

$\gamma_{2j}$ : eigenvector of the environment  $j$  for PCA2;

$\varepsilon_{ij}$ : error that was not explained by both effects.

The biplot graph was then obtained for the effects of the first two principal components (PCA1 vs. PCA2), being the first component associated with the yield ratio, which is due only to the genotype characteristics, and the second component associated with the part of the income due to the G x E interaction (Yan and Holland, 2010; Yan, 2011). Biplots with axis correction were also obtained by means (*Mean* vs. *Stability*), and the biplot PCA1 vs. PCA2 subdivided into macroenvironments with line ranking (*Which Won Where*). Analyses and graphs were generated by the '*GGEbiplotGUT*' package using the R (R Core Team, 2017) computer software.

## RESULTS AND DISCUSSION

The results of the joint deviance analysis, considering all environments, are presented in Table 3. When evaluating lines in the final stages of a breeding program, i.e., value for cultivation and use (VCU), it is essential that they be well conducted and that high experimental accuracy be obtained. Considering the parameters used to measure the experimental accuracy and precision, highest values for accuracy and lowest values for coefficient of variation are desirable. In the present study, as previously mentioned, the accuracy (Resende and Duarte, 2007) and the coefficient of variation (Pimentel-Gomes, 2009) were adopted as a tool to infer on the experimental accuracy and precision. The minimum and maximum values obtained for the accuracy and coefficient of variation were 0.93 and 10.97, respectively, which indicate medium to high experimental accuracy and precision.

Accuracy was affected by the referred trait, as well as by the used analysis strategy. In the individual analyses, lower accuracy was evident. When the joint analysis was adopted, there was an increase in accuracy. This fact can be justified because a greater number of replicates is achieved when the joint analysis is performed, thus guaranteeing greater experimental accuracy and lower error associated to the estimates (Pimentel-Gomes, 2009).

There were significant differences among lines for all evaluated traits. This fact can be explained because the tested lines have genetic differences, due to different genetic background. There were also significant effects of environments on the traits expression. The environmental effect in this study is due to the combination of predictable and unpredictable factors (Allard and Bradshaw, 1964). Both factors are present because the different lines were tested in different sites in the state of Minas Gerais, Brazil, as well as in two crop seasons. The unpredictable factors are particularly important for VCU experiments since the lines must be tested for at least two years in order to quantify this effect and minimize errors in the recommendation of cultivars (Kaster and Farias, 2012).

The genotype x environment interaction was also significant, which allows suggesting that the lines did not show coincident behavior for all the traits in the different evaluation environments. Genotype x environment interaction for soybean crop in the state of Minas Gerais has been reported in the literature (Soares et al., 2015; Gesteira et al., 2015; Silva et al., 2015). There are some alternatives that can be adopted in the occurrence of the genotype x environment interaction to identify the best lines. In this study, it was chosen to identify the most stable lines (Kang and Gauch, 1988).

The adjusted averages of the 39 lines evaluated in the 13 environments for all the characteristics are presented in Table 4. The grain yield averages ranged from 2584.90 kg.ha<sup>-1</sup> to 4515.87 kg.ha<sup>-1</sup> for lines 9 and 18, respectively, with an overall average of 3582.94 kg.ha<sup>-1</sup>. Absolute maturity ranged from 94.64 to 134.39 days also for lines 9 and 18. For the oil content trait, the averages ranged from 19.26 to 23.67% for lines 8 and 6, while the protein content ranged from 34.64 to 41.12% for lines 18 and 8, respectively.

When studying several traits in breeding programs, it is appropriate to quantify the existence of an association among them. In this study, the association among the traits grain yield, absolute maturity, and oil and protein contents present in the grains were obtained (Table 5). The correlation between grain yield and absolute maturity was positive and of high magnitude. Thus, it is expected that to increase the productive potential of cultivars, later lines should be selected (Sediyama, 2015). However, it should also be mentioned that it is possible to obtain early and highly productive cultivars. This fact is possible due to the new strategy adopted by soybean breeding programs in Brazil, focusing on the selection of semi-determined and undetermined growth habit cultivars. In these groups, it is possible to reduce the crop cycle as well as to associate good productive potential (Sediyama, 2015).

Several reports show the success on the use of early soybean cultivars. Gesteira et al. (2015) comment that one of the main advantages on the adoption of these cultivars in the state of Minas Gerais is the possibility of cultivating the second crop ("mid-crop season"). The reduced-cycle

cultivars allow the planting of crop in succession to be carried out in advance, thus ensuring greater utilization of the rainy season, the agricultural areas and higher yields for the producers.

The association between grain yield and protein content was negative and highly significant, being in the same sense of the association between absolute maturity and protein content. It was evident that the lower the absolute maturity, the higher the protein content in the grains, according to the obtained negative correlation and high magnitude. A possible explanation is related to the positive association between oil content and grain yield as well as the negative association between oil and protein contents in the grains (Rincker et al., 2014; Gesteira et al., 2015), which may be related to a lower degradation rate of proteins present in the grains with the decrease of the crop cycle. The higher the productive potential, the longer the crop cycle and higher the oil content in the grains hence lower the protein content (Rincker et al., 2014), which explains the negative association obtained.

The association between the oil and protein contents in soybean has been reported in the literature (Carvalho et al., 2002; Lopes et al., 2002; Carpentieri-Pípolo et al., 2005; Peluzio et al., 2005; Almeida et al., 2010; Nogueira et al., 2012; Rigon et al., 2012) and corroborates with the results obtained in the present study. It can be seen that the strategy adopted by soybean breeding programs in Brazil has been focused on obtaining more productive cultivars. Thus, it is expected an increase in the percentage of oil content in the grains and hence a reduction in the percentage of protein content. However, it should be emphasized that although there is a reduction in the percentage of protein content in the grains, the absolute yield of protein per area may be higher due to the increase in yield.

The index selection is a very useful tool in order to identify and select lines that associate good attributes for several traits. The results for the SMSI are presented in Table 4. The main objective of the index selection was to identify more productive lines, with lower crop cycle and with high oil and protein contents in the grains. It can be observed that the line 27 was more

prominent by SMSI, i.e., associated with early cycle, good yield and grain quality (high oil and protein contents).

The use of index selection in the soybean crop has been reported in the literature. In a study performed by Soares et al. (2015), the authors adopted a multiple index selection and verified that it was efficient in the selection of new soybean cultivars, corroborating the results obtained in this study.

When phenotypic data are available in different environments, an alternative for the identification of more stable cultivars is the use of methods to study the genotype x environment interaction. Several studies considering the genotype x environment interaction in many crop species are presented in the literature (Kendal, 2015; Kendal and Sener, 2015; Dogan et al., 2016; Kendal and Sayar, 2016; Kendal et al., 2016; Kendal and Tekdal, 2016). In the present study, as already mentioned, the GGE Biplot (Yan et al., 2000) was adopted, considering the SMSI. The study of genotype x environment interaction is important because it allows obtaining information about the adaptability and stability of genotypes, especially when it is detected significance in the interaction effect. Analysis by GGE Biplot is an efficient tool for this study in multi-site competition trials (Bhan et al., 2005), and was also considered useful in the study of genotype x environment interaction, being a good tool to generate high quality images and define recommendations for breeders (Kendal and Sayar, 2016). According to Oliveira et al. (2003), the analysis should be performed with the lowest number of axes (principal components) and the highest percentage of explanation possible, because as the number of axes increases, the percentage of noise also increases, which reduces the prediction power of the analysis.

In this context, it is worth pointing out that, in this condition, it is possible to identify stable lines considering not only the trait grain yield, but also important traits, such as absolute maturity and grain quality. Previous studies of GEI and relationship between traits were performed in durum wheat (Kendall, 2015; Kendall and Sener, 2015). The authors found high relationships between grain yield, grain components and quality traits, suggesting that some of them can be used for

indirect selection due to the presented high association. There is only one report in the literature on the interaction study considering a multiple index selection in the soybean crop, however, involving only the traits yield and physiological quality of seeds (Silva, 2016). There were no reports in the literature involving the absolute maturity, yield and grain quality traits.

After the SMSI analysis, the missing values in the double entry tables were predicted for the four characteristics of interest, aiming to recalculate the SMSI in a balancing scenario and to allow studying the genotype x environment interaction. The validation of predictions was also performed by calculating the predictive accuracy based on 1000 resamplings, whose values were 0.71, 0.88, 0.59, and 0.73 for the yield, absolute maturity, and oil and protein contents in the grains, respectively. The obtained values were of high magnitude, proving the efficacy and reliability of predictions and data used to study the genotype x environment interaction.

The analysis of genotype x environment interaction using the GGE Biplot method was efficient, and the first two principal components explained a large part (82.23%) of the observed G + GA effect. This result corroborates the results obtained by Amira et al. (2013), who verified in trials with early soybean lines that the first two principal components explained 86.6% of the interaction effect. Similar results were found in triticale study, where the PCA1 and PCA2 explained 77.50% of the genotype x environment interaction effect (Kendal and Sayar, 2016), while different results were found in spring barley studies, where the PCA1 and PCA2 explained 52.14 and 68.30% of the interaction effect (Dogan et al., 2016; Kendal and Tekdal, 2016).

This method considers the environment effect as irrelevant in the cultivar selection, since the adaptive capacity of genotypes is more important than the environmental conditions (Camargo-Buitrago et al., 2011). In the GGE Biplot graph, the first principal component of the graph (abscissa) is associated with adaptability, also presenting high correlation with average performance, while the second principal component (ordinates) is associated with stability (Yan et al., 2000). It is worth mentioning that the use of SMSI in the GGE Biplot analysis will not indicate



the best performing lines for all the considered traits, but the lines that show averages above the overall averages of these traits in the experiment, associated with phenotypic stability.

Thereby, it was observed that the lines 6, 7, 27 and 31 showed good general adaptability because they showed the highest SMSI values, as well as the superiority in relation to the other lines in the abscissa axis (Figure 1). Among them, the line 7 showed lower stability because it is more distant from the origin of the ordinate axis. Thus, the lines 6, 27 and 31 stood out by their association with good adaptability, stability and averages higher than the other evaluated lines.

Following this same reasoning, environment 13 (Pedrinópolis - MG, 2015/16 crop year) provided the greatest overall adaptability, while environment 5 (Uberaba - MG, 2014/15 crop year) provided the lowest adaptability.

The Figure 2 (*Mean vs. Stability* chart) corrects the positioning of axes with the principal components according to the line averages, thus indicating which the material was with the best performance considering the SMSI (greater value in the abscissa axis), allowing a simultaneous evaluation of its stability by the ordinate axis. The graph provides a combined analysis of adaptability, stability and average performance of the lines, which reinforces the importance of the identification of genotypes with superior performance, besides the safety and predictability in the recommendation of cultivars, seeking to guarantee success with the selection. From this perspective, it can be highlighted that although most of the lines (located in the center of graph) showed good stability and adaptability, their averages were inferior to the averages of lines 27, 6, 31, and 7 mentioned previously, considering the four characteristics simultaneously. It is also possible to highlight the superiority of the line 27 in relation to the others, including the point highlighted on the line as the reference to the ideal environment. Thus, line 27 showed the highest average and therefore better performance and grain quality associated with precocity and stability (Figure 2), reinforcing the result shown in Figure 1.

One of the advantages of GGE Biplot adaptability and stability study is that both axes (PCA1 and PCA2) are in the same unit, which allows the relative comparison of genotypes in each

environment (Yan et al., 2000). A good tool to compare the relative performance of genotypes in environments is the *Which Won Where* chart, which subdivides genotypes and environments into regions called mega-environments, which are delimited by red lines and a polygon. The genotypes and environments contained in each mega-environment have specific adaptability among themselves, and the lines that form the polygon vertices showed the best performance within each region and therefore are the most promising in the environments contained in the referred region.

The obtained graph allows visualizing the formation of a mega-environment delimited by the line 27 at its extreme vertex (Figure 3), also encompassing several other lines with a performance superior to the general average of index. The environments 9, 7 and 4 are also contained in this mega-environment, and it can be suggested that the lines contained in this region showed good specific adaptability for the mentioned environments. However, due to the greater proximity among the lines 27, 6, and 31 with these environments and the fact that both axes present the same unit, it can also be commented that the specific adaptability among these lines and the mentioned environments is greater in relation to the other lines contained in this mega-environment.

Based on this analysis, it can be suggested that the lines 27, 6, and 31 showed superior performance in a larger number of environments and hence contributed less to the interaction in relation to the other evaluated lines. These lines associated specific adaptability and good stability to good performance for the traits contemplated by the SMSI.

## CONCLUSION

Lines 6, 27, and 31 show phenotypic stability associated with good performance in grain yield and quality, considering the SMSI.

There is a negative and high magnitude association between the absolute maturity and protein content in the grains, indicating that it is possible to obtain early lines with high protein content in the grains.

The GGE Biplot method is efficient to quantify the genotype x environment interaction when the SMSI is adopted, constituting an important tool for selection and recommendation of cultivars.

## REFERENCES

- Allard, R. W. and Bradshaw, A. D. 1964. Implications of genotype-environmental interactions in applied plant breeding. *Crop Science*. 4:503-508.
- Almeida, L.A. and Kiihl, R.A.S. 1998. Melhoramento da soja no Brasil - desafios e Perspectivas. In: *Soja: Tecnologia da Produção*. Gil. M. S. Câmara (ed.). p. 40-54.
- Almeida, R. D., Peluzio, J. M. and Afferi, F. S. 2010. Correlações fenotípicas, genotípicas e ambientais em soja cultivada sob condições várzea irrigada, sul do Tocantins. *Bioscience Journal*. 26:95-99.
- Amira, J.O., Ojo, D.K., Ariyo, O.J., Oduwaye, O.A. and Ayo-Vaughan, M.A. 2013. Relative discriminating powers of GGE and AMMI models in the selection of tropical soybean genotypes. *African Crop Science Journal*. 21:67-73.
- Barros, H. B., Sediya, T., Fidelis, R. R. and Capone, A. 2012. Adaptabilidade e estabilidade de genótipos de soja por meio de métodos uni e multivariado. *Journal of Biotechnology and Biodiversity*. 3:49-58.
- Bhan, M.K., Pal, S., Rao, B.L., Dhar, A.K. and Kang, M.S. 2005. GGE biplot analysis of oil yield in lemon-grass (*Cymbopogon* spp.). *Journal of New Seeds*. 7:127-139.
- Bueno, R. D., Borges, L. L., Arruda, K. M. A., Behring, L. L., Barros, E. G. and Moreira, M. A. 2013. Genetic parameters and genotype x environment interaction for productivity, oil and protein content in soybean. *African Journal of Agricultural Research*. 8:4853-4859.
- Camargo-Buitrago, I., Intire, E.M. and Córdon-Mendoza, R. 2011. Identificación de mega-ambientes para potenciar el uso de progênes superiores de arroz en Panamá. *Pesquisa Agropecuária Brasileira*. 46:1601-1069.
- Carpentieri-Pípolo, V., Gastaldi, L. F. and Pipolo, A. E. 2005. Correlações fenotípicas entre caracteres quantitativos em soja. *Semina: Ciências Agrárias*. 26:11-16.
- Carvalho, C. G. P., Arias, C. A. A., Toledo, J. F. F., Oliveira, M. F. and Vello, N. A. 2002. Correlações e análise de trilha em linhagens de soja semeadas em diferentes épocas. *Pesquisa Agropecuária Brasileira*. 37:311-320.
- Carvalho, C. G. P., Arias, C. A. A., Toledo, J. F. F., Almeida, L. A., Kiihl, R. A. S. and Oliveira, M. F. 2002. Interação genótipo x ambiente no desempenho produtivo da soja no Paraná. *Pesquisa Agropecuária Brasileira*. 37:989-1000.
- Companhia Nacional De Abastecimento (CONAB). 2016. Acompanhamento da safra brasileira de grãos – Safra 2015/16. 3:1-184.
- Cruz, C. D., Regazzi, A. J. and Carneiro, P. C. S. 2012. Modelos Biométricos Aplicados ao Melhoramento Genético. 4th edition. Ed. UFV. 514p.
- Dogan, Y., Kendal, E. and Oral, E. 2016. Identifying of relationship between traits and grain yield in spring barley by GGE Biplot analysis. *Agriculture & Forestry*. 62:239-252.
- Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). 2010. Boletim de Pesquisa e Desenvolvimento: Índices pluviométricos em Minas Gerais. 88 p.
- Empresa Brasileira De Pesquisa Agropecuária (EMBRAPA). 2013. Tecnologias de produção de soja – Região Central do Brasil 2014. Embrapa Soja. 265p.
- Gesteira, G.S., Zambiazzi, E.V., Bruzi, A.T., Soares, I.O., Rezende, P.M. and Silva, K.B. 2015. Seleção fenotípica de cultivares de soja precoce para a região Sul de Minas Gerais. *Revista Agrogeoambiental*. 7:79-88.
- Kang, M. S. and Gauch Jr., H. G. 1996. Genotype by environment interaction. CRC Press. 416 p.
- Kaster, M. and Farias, J.R.B. 2012. Regionalização dos testes de valor de cultivo e uso e da indicação de cultivares de soja: terceira aproximação. Embrapa Soja.
- Kendal, E. 2015. Relationship between Chlorophyll and other Features in Durum Wheat (*Triticum turgidum* L. var. durum) Using SPAD and Biplot Analyses. *J. Agr. Sci. Tech*. 17:1873-1886.

- Kendal, E., Sayar, M. S. 2016. The stability of some spring triticale genotypes using Biplot analysis. *The Journal of Animal & Plant Sciences*. 26:754-765.
- Kendal, E., Sayar, M.S., Tekdal, S., Aktas, H. and Karaman, M. 2016. Assessment of the impact of ecological factors on yield and quality parameters in triticale using GGE Biplot And AMMI analysis. *Pak. J. Bot.* 48:1903-1913.
- Kendal, E., Sener, O. 2015. Examination of genotype  $\times$  environment interactions by GGE biplot analysis in spring durum wheat. *Indian J. Genet.* 75:341-348.
- Kendal, E., Tekdal, S. 2016. Application of AMMI model for evaluation spring barley genotypes in multi-environment trials. *Bangladesh J. Bot.* 45:613-620.
- Lima, W. F., Pípelo, A. E., Moreira, J. U. V., Carvalho, C. G. P., Prete, C. E. C., Arias, C. A. A., Oliveira, M. F., Souza, G. E. and Toledo, J. F. F. 2008. Interação genótipo-ambiente da soja convencional e transgênica resistente a glifosato, no Estado do Paraná, Brasil. *Pesquisa Agropecuária Brasileira*. 43:729-736.
- Lopes, A. C. A., Vello, N. A., Pandini, F., Rocha, M. M. and Tsutsumi, C. Y. 2002. Variabilidade e correlações entre caracteres em cruzamentos de soja. *Scientia Agricola*. 59:341-348.
- Marques, M. C., Hamawaki, O. T., Sediya, T., Bueno, M. R., Cruz, C. D., Reis, M.S. and Nogueira, A.P.O. 2011. Adaptabilidade e estabilidade de genótipos de soja em diferentes épocas de semeadura. *Bioscience Journal*. 27:59-69.
- Mendes, F.F., Ramalho, M.A.P. and Abreu, A.F.B. 2009. Selection index for choosing segregating populations in common bean. *Pesquisa Agropecuária Brasileira*. 44:1312-1318.
- Nogueira, A. P. O., Sediya, T., Sousa, L. B., Hamawaki, O. T., Cruz, C. D., Pereira, D. G. and Matsuo, E. 2012. Análise de trilha e correlações entre caracteres em soja cultivada em duas épocas de semeadura. *Bioscience Journal*. 28:877-888.
- Oliveira, B.A., Duarte, J.B. and Pinheiro, J.B. 2003. Emprego da análise AMMI na avaliação da estabilidade produtiva em soja. *Pesquisa Agropecuária Brasileira*. 38:357-364.
- Peluzio, J. M., Almeida, R. D., Fidelis, R. R., Júnior, D. A., Brito, E. L. and Francisco, E. R. 2005. Correlações entre caracteres de soja, em Gurupi, Tocantins. *Revista Ceres*. 52:779-786.
- Pimentel-Gomes, F. 2009. Curso de estatística experimental. 15th edition, Ed. FEALQ, 451p.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (<https://www.R-project.org/>).
- Resende, M.D.V. and Duarte, J.B. 2007. Precisão e controle de qualidade em experimentos de avaliação de cultivares. *Pesquisa Agropecuária Tropical*. 3:182-194.
- Rigon, J. P. G., Capuani, S., Brito Neto, J. F., Rosa, G. M., Wastowski, A. D. and Rigon, C. A. G. 2012. Dissimilaridade genética e análise de trilha de cultivares de soja avaliada por meio de descritores quantitativos. *Revista Ceres*. 59:233-240.
- Rincker, K., Nelson, R., Specht, J., Sleper, D., Cary, T., Ciazio, S.R., Casteel, S., Conley, S., Chen, P., Davis, V., Fox, C., Graef, G., Godsey, C., Holshouser, D., Jiang, G., Kantartzi, S.K., Kenworthy, W., Lee, C., Mian, R., McHale, L., Naeve, S., Orf, J., Poysa, V., Schapaugh, W., Shannon, G., Uniatowski, R., Wang, D. and Diers, B. 2014. Genetic improvement of US soybean in maturity groups II, III, and IV. *Crop Science*. 54:1419-1432.
- Scott, A. J. and Knott, M. A. 1974. Cluster analysis method for grouping means in the analysis of variance. *Biometrics*. 30:507-512.
- Sediya, T. 2015. Melhoramento genético da soja. Londrina: Ed. Mecenias, v.1, 352p.
- Silva, K. B., Bruzi, A. T., Zuffo, A. M., Zambiazzi, E. V., Soares, I. O., Rezende, P. M., Fronza, V., Vilela, G. D. L., Botelho, F. B. S., Teixeira, C.M. and Coelho, M.A.O. 2015. Adaptability and phenotypic stability of soybean cultivars for grain yield and oil content. *Genetics and Molecular Research*. 15:1-15.
- Silva, K.B. 2016. Plasticidade fenotípica e análise de QTL para qualidade de sementes em soja. 108 p. Thesis (PhD in Agronomy/Plant Sciences) - Universidade Federal de Lavras.

- Soares, I.O., Rezende, P.M., Bruzi, A.T., Zambiazzi, E.V., Felisberto, G., Zuffo, A.M., Silva, K.B. and Gwinner, R. 2015. Adaptability of soybean cultivars in different crop years. *Genetics and Molecular Research*. 14:8995-9003.
- Vasconcelos, E.S., Reis, M.S., Cruz, C.D., Sediya, T. and Scapim, C. A. 2010. Adaptability and stability of semilate and late maturing soybean genotypes in Minas Gerais state. *Acta Scientiarum*. 32:411-415.
- Yan, W., Hunt, L.A., Sheng, Q. and Szlavniks, Z. 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science*. 40:597-605.
- Yan, W. and Holland, J.B. 2010. A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica*. 171:355-369.
- Yan, W. 2011. GGE Biplot vs. AMMI Graphs for Genotype-by-Environment Data Analysis. *Journal of the Indian Society of Agricultural Statistics*. 65:181-193.
- Yan, W. 2013. Biplot Analysis of Incomplete Two-Way Data. *Crop Science*. 53:48-57.

## SUPPLEMENTARY MATERIALS

**Table 1** – Crop seasons, trial sites, latitude and longitude (geographical), annual rainfall and correspondent environments.

Crop Seasons	City	Latitude	Longitude	Annual Rainfall (mm)	Environment
2014/15	Conceição das Alagoas	-19.9147	-48.3883	1485.6	1
2014/15	Iraí de Minas	-22.0989	-46.1897	1309.0	2
2014/15	Presidente Olegário	-18.4178	-46.4181	1504.0	3
2014/15	Sacramento	-19.8653	-47.4400	1630.2	4
2014/15	Uberaba	-19.7667	-47.9333	1528.5	5
2014/15	Muzambinho	-21.3758	-46.5256	1488.2	6
2015/16	Conceição das Alagoas	-19.9147	-48.3883	1485.6	7
2015/16	Presidente Olegário	-18.4178	-46.4181	1504.0	8
2015/16	Sacramento	-19.8653	-47.4400	1630.2	9
2015/16	Uberaba	-19.7667	-47.9333	1528.5	10
2015/16	Água Comprida	-20.0564	-48.1089	1409.0	11
2015/16	Buritzeiro	-17.3511	-44.9622	1086.9	12
2015/16	Pedrinópolis	-19.2278	-47.4622	1605.6	13

**Table 2** – Treatments and their corresponding cultivars (1-5) and inbred lines (6-39).

Treatment	Line	Treatment	Line
1	NA 5909 RR	21	BRRY34-0901
2	BMX Potência RR	22	RRMG11-55821
3	BMX Desafio RR	23	BRRR12-67015
4	Anta 82 RR	24	BRRR12-67205
5	NA 7337 RR	25	BRRY34-1172
6	BR09-1554	26	BRRY34-1175
7	RRMG09-99717	27	BRRY45-10190
8	RRMG11-53504	28	BRRY45-10378
9	RRMG11-54115	29	BRRY45-10473
10	RRMG11-56011	30	BRRY45-10649
11	RRMG11-57119	31	BRRY45-10729
12	BRRR12-67204	32	BRY34-1245
13	BRRR12-70304	33	RRMG12-69204
14	BRY23-0068	34	RRMG12-78605
15	BRY23-0080	35	RRMG13-85603
16	BRRY34-0591	36	RRMG13-85908
17	BRY23-0035	37	RRMG13-87803
18	BRY23-0089	38	RRMG13-88501
19	BRY23-0062	39	RRMG13-97816
20	BRRY34-0784		

**Table 3** – Summary of joint deviance analysis, considering all environments, for all evaluated traits.

Sources	d.f.	Grain Yield		Absolute Maturity		Oil		Protein	
Random effects:	-	Var.	p-value	Var.	p-value	Var.	p-value	Var.	p-value
L x E	456	22.57	<2e-16 **	16.36	<2e-16 **	0.20	3e-04 **	0.59	4e-16 **
Rep(E)	39	6.12	9e-12 **	0.61	2e-05 **	-	-	-	-
Environments (E)	12	97.24	<2e-16 **	57.40	<2e-16 **	0.15	5e-10 **	0.27	1e-08 **
Residuals	1482 (507)	56.48	-	10.35	-	0.70	-	0.67	-
Fixed effects:	-	F	Pr(>F)	F	Pr(>F)	F	Pr(>F)	F	Pr(>F)
Lines (L)	39	7.92	0.00	34.18	0.00	8.30	0.00	10.34	0.00
CV (%)	-	10.97		8.27		3.80		3.42	
Accuracy	-	0.93		0.99		0.94		0.95	
Averages	-	3582.94		114.90		21.98		37.34	

d.f.: degrees of freedom (residuals for oil and protein contents in parenthesis); Var.: Variance; CV: Coefficient of variation; \*\* significant at 1% probability.



**Table 4** – Adjusted averages for evaluated traits and for standardized multiple selection index (SMSI) in all inbred lines, obtained from the joint analysis.

Line	Grain Yield (kg.ha <sup>-1</sup> )	Absolute Maturity (days)	Oil (%)	Protein (%)	SMSI
1	3793.58 g	109.42 s	23.33 c	35.77 p	8.35
2	3670.90 i	115.00 m	22.39 h	36.99 l	7.30
3	4202.25 c	124.92 e	22.54 g	36.70 m	7.20
4	3782.70 g	123.50 g	21.90 l	37.14 k	6.27
5	4000.09 d	129.64 b	21.73 n	36.25 n	4.77
6	3775.20 g	109.81 r	23.67 a	37.12 k	9.04
7	3582.05 k	109.19 t	22.69 f	37.60 i	8.48
8	2807.65 t	98.14 x	19.26 u	41.12 a	6.69
9	2584.90 u	94.64 z	21.96 l	38.13 f	7.07
10	3444.50 n	103.75 w	21.86 m	37.51 j	8.02
11	3171.87 q	105.89 v	22.10 k	37.74 h	7.76
12	3544.23 l	111.25 p	22.94 d	37.68 h	8.33
13	3684.32 i	112.19 o	22.12 k	37.77 h	7.40
14	4439.92 b	129.85 b	21.17 q	37.02 l	5.48
15	3963.82 d	126.85 d	21.78 m	35.13 q	4.91
16	3480.83 m	120.39 h	22.34 i	36.11 o	5.84
17	3869.05 f	128.85 c	22.43 h	36.06 o	6.40
18	4515.87 a	134.39 a	22.02 k	34.64 r	5.56
19	3774.25 g	126.94 d	22.15 k	36.20 n	6.18
20	3907.04 e	124.19 f	21.70 n	37.64 i	6.98
21	3626.20 j	120.10 i	22.47 g	35.19 q	5.95
22	2860.50 s	95.60 y	21.20 q	38.10 f	7.73
23	3753.86 g	111.41 p	21.42 o	38.06 f	7.75
24	3581.81 k	109.98 r	22.17 k	37.77 h	8.06
25	3184.95 q	115.98 k	21.30 p	38.87 c	6.27
26	3543.88 l	116.14 k	20.88 s	38.75 d	6.52
27	3534.24 l	109.73 r	23.56 b	37.62 i	9.30
28	3313.81 o	112.68 n	22.34 i	38.16 f	7.45
29	3581.94 k	117.19 j	22.25 j	37.64 i	6.89
30	3035.19 r	116.01 k	22.87 d	36.30 n	5.26
31	3647.19 j	107.88 u	22.22 j	38.41 e	9.14
32	3206.58 q	128.70 c	22.10 k	35.78 p	2.58
33	3707.02 h	110.73 q	21.84 m	37.94 g	7.96
34	3619.55 j	112.84 n	20.98 r	38.35 e	7.03
35	3641.83 j	109.98 r	20.63 t	39.45 b	7.91
36	3506.70 m	115.73 l	21.32 p	38.38 e	6.74
37	3715.23 h	109.13 t	22.79 e	35.79 p	7.44
38	3410.08 n	109.91 r	21.65 n	37.85 h	7.19
39	3269.19 p	112.38 o	21.07 r	37.44 j	5.52
Averages	3582.94	114.90	21.98	37.34	6.94

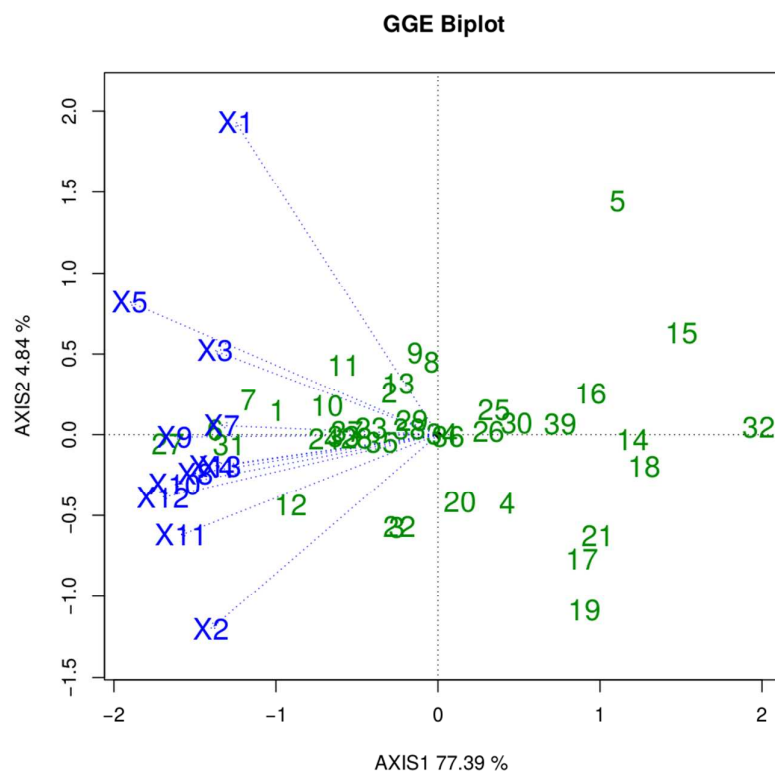
Averages followed by the same letter belong to the same group, by Scott and Knott (1974) at 5% probability.

**Table 5** – Pearson estimated correlations between all evaluated traits, considering adjusted averages shown in Table 4.

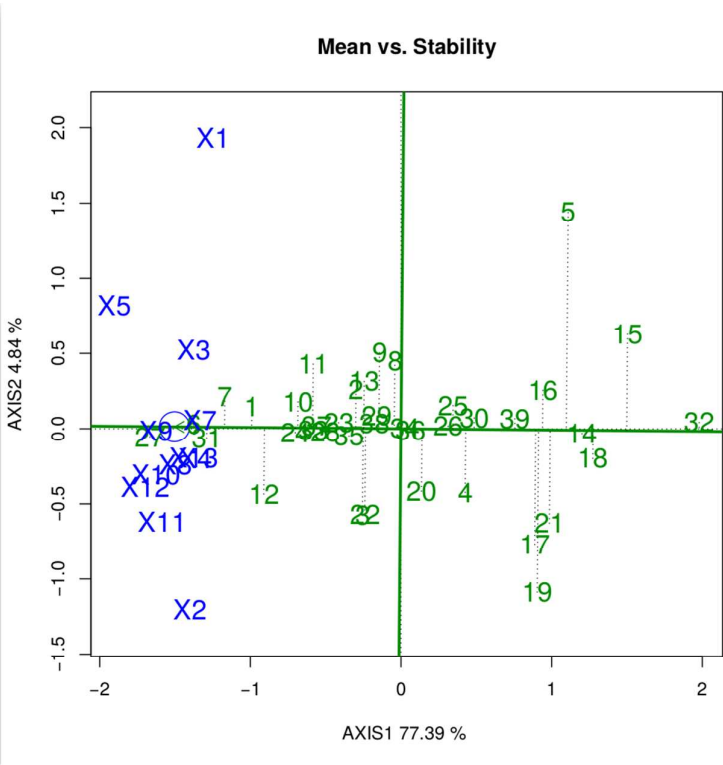
	AM	Oil	Protein
Grain Yield	0.7204 **	0.2181	-0.4998 **
AM		0.1151	-0.6306 **
Oil			-0.5884 **

AM: absolute maturity; \*\* significant at 1% probability.

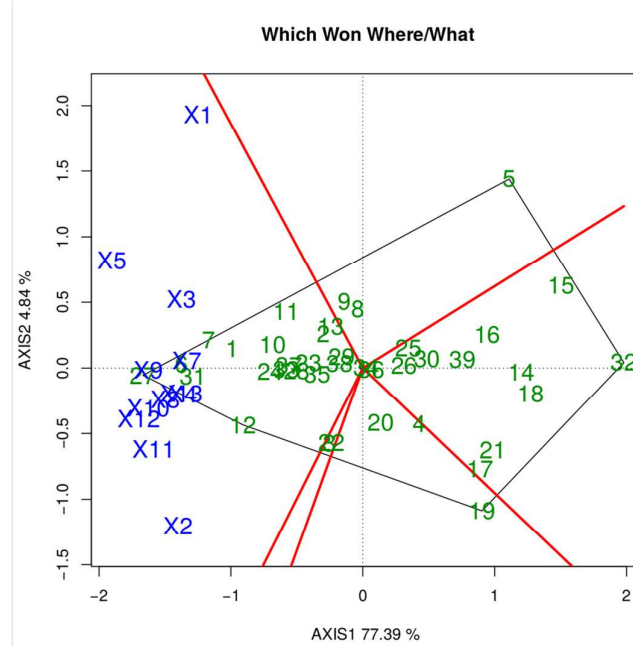
**Figure 1** - Biplot with the effects of the first two principal components (PCA1 vs. PCA2) for the standardized multiple selection index (SMSI) of the 39 lines evaluated in 12 environments. The genotypes are represented by the blue coloration, while the environments are represented by the green coloration.



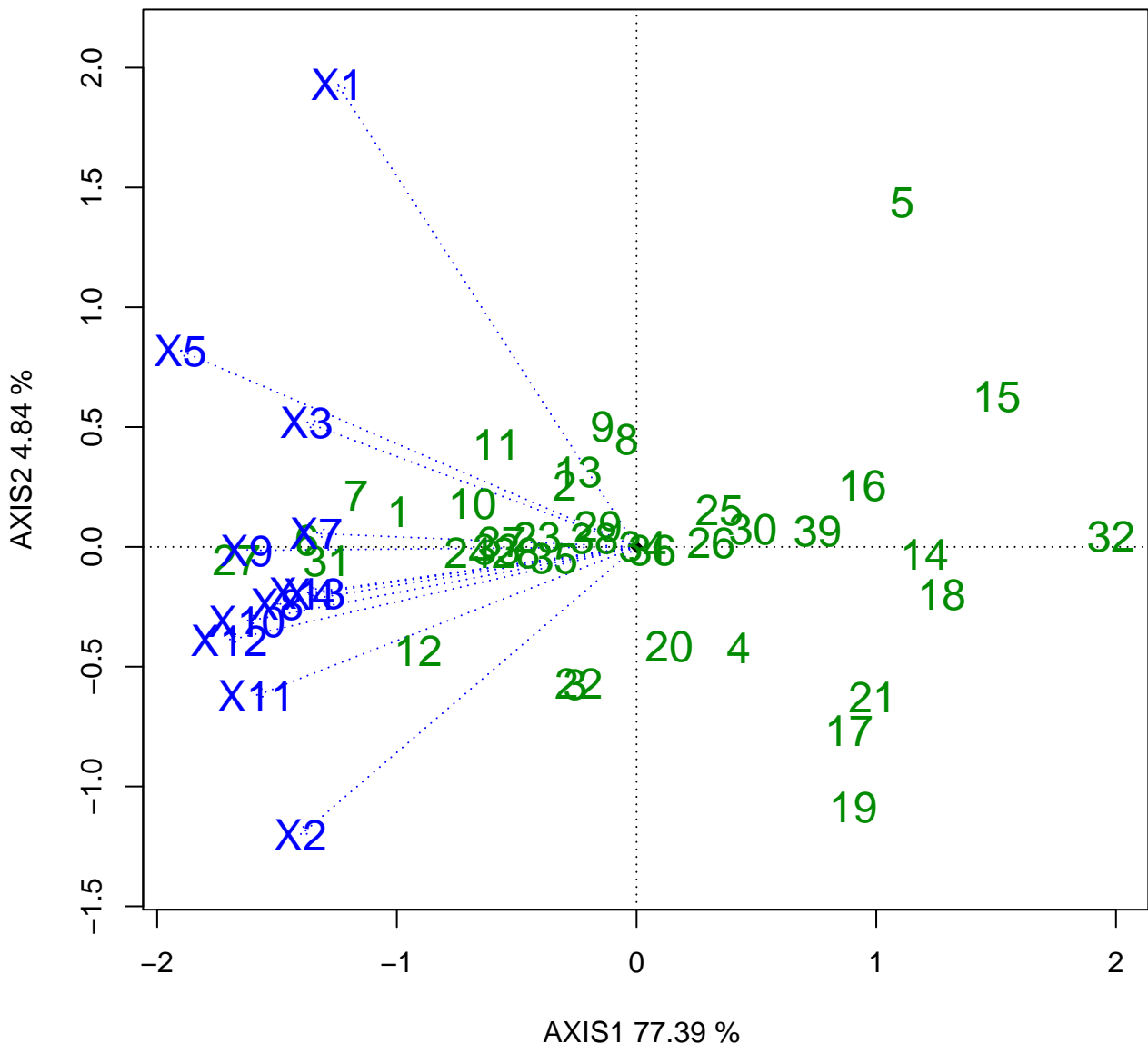
**Figure 2** - Biplot of the principal components PCA1 vs. PCA2 in the main axes, with second positioning of axes in relation to the genotype average for the standardized multiple selection index (SMSI) of the 39 evaluated lines in 12 environments. The genotypes are represented by the blue coloration, while the environments are represented by the green coloration.



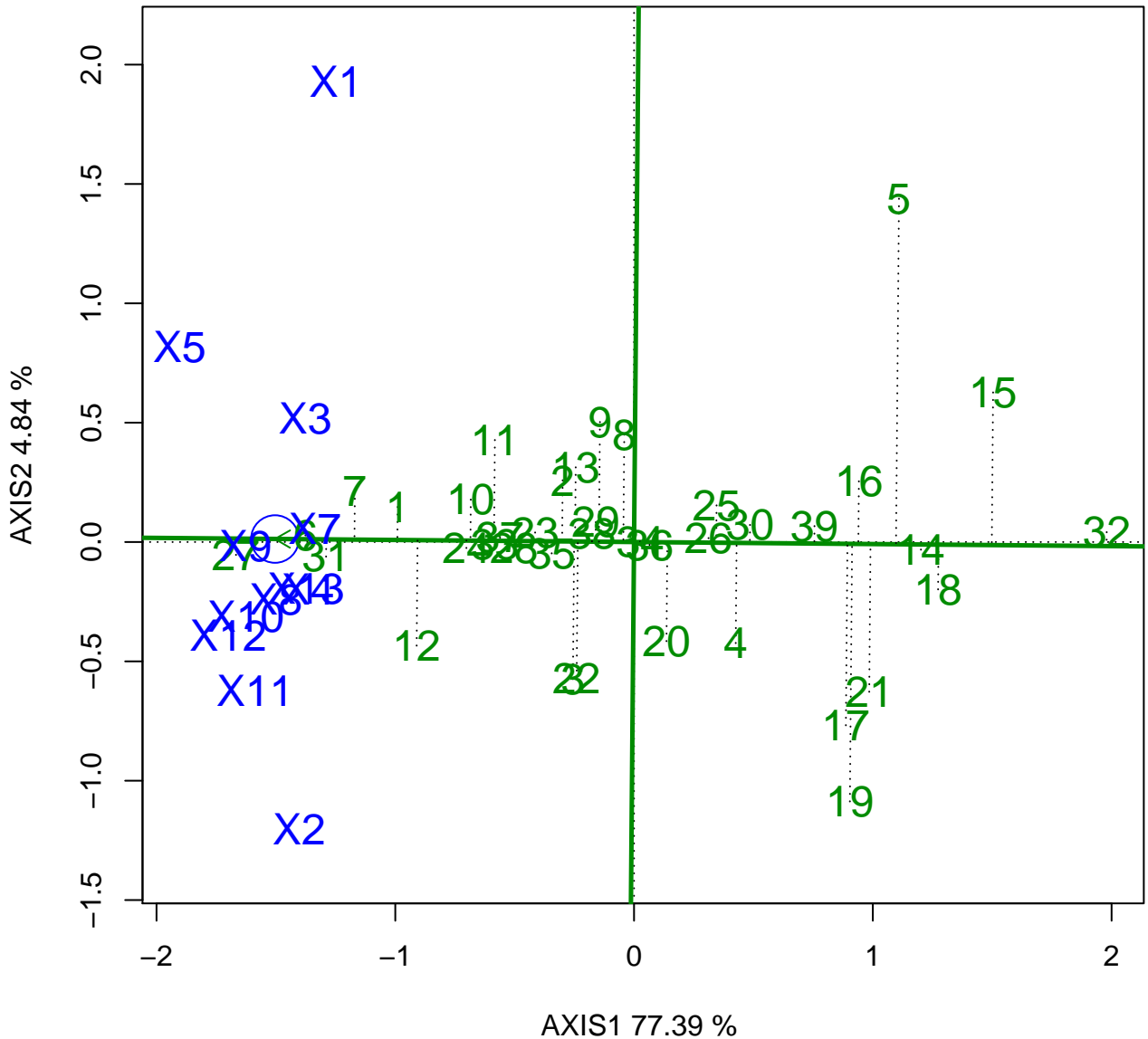
**Figure 3** - Biplot of the principal components PCA1 vs. PCA2 in the main axes, with subdivision of the macro-environments and ranking considering the SMSI of the 39 lines evaluated in the 12 environments. The genotypes are represented by the blue coloration, while the environments are represented by the green coloration.



GGE Biplot



# Mean vs. Stability



Which Won Where/What

