RESEARCH ARTICLE

Correlations and canonical variables applied to the distinction of soybean cultivars in a tropical environment

Luiz Leonardo Ferreira¹, Ângelo José da Silva², Ivan Ricardo Carvalho²,* Marilaine de Sá Fernades¹, Francine Lautenchleger³ and Murilo Vieira Loro⁴

¹Universidade em Mineiros, Goiás, GO, Brazil. ²Regional University of the Northwest of the State of Rio Grande do Sul, Rua do Comércio, 3000, Universitário, Ijuí, RS, Brazil, CEP 98700-000. ³Universidade Estadual do Centro Oeste, Guarapuava, PR, Brazil, CEP 85015-430. ⁴Federal University of Santa Maria, Santa Maria, RS, Brazil. *Corresponding author, E-mail: carvalho.irc@gmail.com

ABSTRACT

The objective of this study was to evaluate the performance of soybean cultivars through their correlations and canonical variables in a tropical environment. This study was conducted in the municipality of Mineiros, GO, Brazil. The experimental design used was randomized blocks with four replications using 10 soybean cultivars (Bônus, Desafio, Flecha, Foco, ICS7019, M5917, M7110, Power, ST721 and ST797). During the conduct of the experiment, pest control was carried according to good practices and integrated management. At the end of the cycle of each cultivar, 10 plants were collected at random and then the agronomic attributes were taken. The data obtained were submitted to the assumptions of the statistical model, verifying the normality and homogeneity of the residual variances, as well as the additivity of the model. Univariate and multivariate models were used. The analysis were performed on the Rbio software using the R platform in addition to the Software Genes. According to the summary of analysis of variance, it was observed that all cultivars differed for all traits. The soybean cultivars Flecha and M5917 showed the highest yields in a tropical environment. The cultivars showing a strong correlation between the number of grains per plant and yield. And the canonical correlations tools were efficient and complementary in the data analysis.

Keywords: Competition test, Glycine max, plant genetics, biometrics, soybean yield, cultivar adaptation.
INTRODUCTION

Soybean (*Glycine max*) is the most widely cultivated oilseed in the world and the basis for the food of several peoples, representing an important source of raw material for industry, human and animal nutrition, being one of the most important commodities for generating favorable trade balance (Ribeiro et al., 2016).

Since its domestication, people have been interested in its large-scale production with a view to making profits and taking advantage of its great energetic, nutritious and commercial potential. Soybean, despite not being considered as a staple food, is one of the most important crops in the world, mainly for the supply of protein and vegetable oil. Furthermore, according to Loro et al. (2021), crop grains are sources of macronutrients such as nitrogen and phosphate with 51.9 and 3.7 g kg$^{-1}$ of grains, respectively, as well as micronutrients such as boron, iron, manganese, zinc and molybdenum with concentrations up to 108.7 mg kg$^{-1}$. According to Oda, Sediyama and Cruz (2022), soybean farming had a rapid expansion in Brazil, due to its economic value, nutritional and thanks to its development, through genetic improvement of new cultivars more adapted to the country's conditions.

As the development of soybean culture has become important for the Brazilian economy, this sector is being one of the main pillars responsible for the consolidation and stabilization of the national economy. In the 2020/2021 harvest, Brazil presented an area of 38,529.0 thousand hectares of soybean plantation, and a production of 135,978.3 thousand tons (Companhia Nacional de Abastecimento [CONAB], 2021; Véras, Matsuo, Dias-Pereira, Ferreira, & Rocha, 2021).

According to Campos, Costa, Almeida, & Simon (2016), in order to obtain high yield in soybean culture, it is necessary to have, among other factors, availability of water and nutrients and effective control of pests and diseases, combined with the use of plants with high productive potential and adapted to local conditions (Gomes, Pimenta, Amaral, Rodrigues, & Borem, 2021; Pagliarini et al., 2021). Each year breeding companies create new cultivars to meet the demands of farmers who face the need to seek higher levels of yield (Borges et al., 2018).

For soybean seed producers, the main objective is to achieve greater crop profitability, reaching high yield and quality per hectare harvested (Marcos-Filho, 2015). To this end, several techniques for integrated management of insect pests, diseases and weeds are added, as well as the preparation of suitable soil, the use of good quality seeds, associated with the choice of genetic materials adapted to the region, the spatial arrangement of plants in the area (Cruz, Sena-Junior, Santos, Lunezzo, & Machado, 2016) and the sowing time (Follmann et al., 2016). With the launch of Intacta RR2 PRO technology and INOX technology, which gives resistance to the active glyphosate, main defoliating caterpillars and the Soybean rust disease (*Phakopsora pachyrhizie*), it is expected to add more advantages to the no-tillage system in the region, as it will contribute to reduce the traffic with sprays, consequently there will be less soil compaction and a significant cost reduction for producers (Cordeiro-Júnior et al., 2017).

Farmers are increasingly looking for the best positioning of each cultivar in order to achieve maximum production, consequently increasing the profit margin in marketing. And so it is clear that in an agricultural company there are different production environments, and it is common knowledge that the properties have differences in edaphoclimatic conditions, which directly influences the cost of inputs, such as the acquisition of seed that must express all its productive potential, associated with the management used, fertility, plant arrangement, pest and disease control, and favorable climatic conditions.
The search for genetic material resistant or tolerant to pests, diseases and herbicides always associated with a high productive potential is what large companies seek, which annually launch two to three new materials on the market. The study of the adaptation of cultivars to different microclimates becomes important for the knowledge of the best productive material and its behavior in the face of environmental variations (Gaviraghi, Pellegrin, Werner, Bellé, & Basso, 2018). Furthermore, the use of strategies such as estimating variance components and genetic parameters of aspects of agronomic importance contribute to genotype selection based on the agronomic ideotype and thus promote better crop performance (Barbosa et al., 2021).

In this way, the importance of soybean for Brazil is great, and in order to obtain high levels of yield, the challenge is to plant a cultivar that has good quality, and that best adapts to the region, so given the above, this study aimed evaluate the performance of soybean cultivars through their correlations and canonical variables in a tropical environment.

MATERIAL AND METHODS

The study was conducted at the Luís Eduardo de Oliveira Salles Experimental Farm, belonging to the Mineiros University Center - UNIFIMES, rural area of the municipality of Mineiros, GO, Brazil. Geographically it is at 17º 58' S latitude and 45º 22' W longitude and approximately 800 m altitude. Average temperature of 22.7 °C and average annual rainfall of 1695 mm, occurring mainly in spring and summer. The experimental area is classified as Aw type (hot to dry) (Köppen & Geiger, 1936).

The results of chemical analyzes of soil samples in the 0-20 cm layer collected in the experiment area were: hydrogen potential 5.7; calcium 3, magnesium 0.8, aluminum 0.2, hydrogen + aluminum 2, cation exchange capacity 5.9, in cmolc dm⁻³; potassium 53, phosphorus 59, sulfur 1.7, boron 0.2, copper 1.4, iron 51, manganese 23, zinc 8.3, sodium 0.5, in mg dm⁻³; clay 223, silt 50, sand 728, organic matter 20, and organic carbon 12, in g dm⁻³. The data were taken according to the methodology of (Empresa Brasileira de Pesquisa Agropecuária [EMBRAPA], 2009). The soil was classified as a Quartzarenic Neosol (Entisol) of sandy texture (Empresa Brasileira de Pesquisa Agropecuária [EMBRAPA], 2013).

The experimental design used was randomized blocks with four replications using 10 soybean cultivars (Bônus, Desafio, Flecha, Foco, ICS7019, M5917, M7110, Power, ST721 and ST797). The experimental plots were composed of 4 lines of 5 meters long, with a spacing of 0.45 m between the lines, the useful area of the plot will be 9 m², with a density of 15 seeds per meter of furrow. The main morpho-agronomic traits of soybean cultivars were described in (Table 1).

Before planting, pre-planting desiccation was performed (Cobucci, Stefano, & Kluthcouski, J.). It was used 450 kg ha⁻¹ of fertilizer 05-25-15 applied in the furrow and in a single dose next to the seeding. Sowing was carried out on November 8, 2018 (Ferreira, Amaral, Silva, Curvelo, & Pereira, 2019). During the conduct of the experiment, the control of pests, diseases and weeds were carried out as they became necessary, respecting good practices and integrated management (Quintela, 2001).

At the end of the cycle of each cultivar, 10 plants were collected at random from the useful area of the experimental plot, thus, the following agronomic attributes of the cultivars were evaluated: POG: pods with one grain in units; PTWG: pods with two grains in units; PTHG: pods with three grains in units; PFG: pods with four grains in units; NGP: number of grains per pod in units; NPP: number of pods per plant in
units; GP: number of grains per plant in units; and YI: yield in sc ha\(^{-1}\).

Table 1. Morpho-agronomic traits of the soybean cultivars analyzed. Mineiros, GO, Brazil. UNIFIMES, 2020.

<table>
<thead>
<tr>
<th>Cultivar name</th>
<th>Thousand seed mass (g)</th>
<th>Company</th>
<th>Maturation group</th>
<th>Growth</th>
<th>Cycle (days after emergence)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bônus 8579 RSF IPRO</td>
<td>190</td>
<td>Bramax</td>
<td>7.9</td>
<td>Indeterminate</td>
<td>105 to 122</td>
</tr>
<tr>
<td>Desafio RR 8473 RSF</td>
<td>180</td>
<td>Bramax</td>
<td>7.4</td>
<td>Indeterminate</td>
<td>105 to 115</td>
</tr>
<tr>
<td>Flecha 6266RSF IPRO</td>
<td>190</td>
<td>Bramax</td>
<td>6.6</td>
<td>Indeterminate</td>
<td>95 to 105</td>
</tr>
<tr>
<td>Fogo 7477 RSF IPRO</td>
<td>175</td>
<td>Bramax</td>
<td>7.4</td>
<td>Indeterminate</td>
<td>110 to 115</td>
</tr>
<tr>
<td>ICS 7019 RR</td>
<td>170</td>
<td>Intelicrops</td>
<td>7.0</td>
<td>Indeterminate</td>
<td>110 to 112</td>
</tr>
<tr>
<td>M 5917 IPRO</td>
<td>185</td>
<td>Monsoy</td>
<td>5.9</td>
<td>Indeterminate</td>
<td>95 to 105</td>
</tr>
<tr>
<td>M 7110 IPRO</td>
<td>175</td>
<td>Monsoy</td>
<td>6.8</td>
<td>Indeterminate</td>
<td>95 to 105</td>
</tr>
<tr>
<td>BMX Power IPRO</td>
<td>170</td>
<td>Bramax</td>
<td>7.3</td>
<td>Indeterminate</td>
<td>105 to 115</td>
</tr>
<tr>
<td>ST 721 IPRO</td>
<td>165</td>
<td>Monsanto</td>
<td>7.2</td>
<td>Indeterminate</td>
<td>105 to 110</td>
</tr>
<tr>
<td>ST 797 IPRO</td>
<td>140</td>
<td>Monsanto</td>
<td>7.9</td>
<td>Indeterminate</td>
<td>115 to 120</td>
</tr>
</tbody>
</table>

The data obtained were submitted to the assumptions of the statistical model, verifying the normality and homogeneity of the residual variances, as well as the additivity of the model. Afterwards, the analysis of variance and application of the Scott-Knott averaging test was performed, at 5% probability. Subsequently, the variables were subjected to Pearson’s linear correlation in order to understand the association trend, with its significance based on a 5% probability by the t test. Canonical correlations were estimated between group 1 (GP and YI) and group 2 (POG, PTWG, PTHG, PFG, NGP and NPP), with significance between the groups of traits assessed based on the chi-square statistic. After the genetic dissimilarity was carried out by the Mahalanobis algorithm, where the residual matrix was weighted, the distance dendrogram was constructed using the UPGMA cluster, then the biplot canonical variables method was used, where it was possible to visualize the general variability of the experiment and the multivariate trends. The analysis were performed on the Rbio software using the R platform (Bhering, 2017), in addition to the Software Genes (Cruz, 2016).

RESULTS AND DISCUSSION

According to the summary of analysis of variance, it was observed that all cultivars differed for all traits analyzed (p < 0.01). Reliability was observed in the values of CV’s with low values (CV < 10) for NGP: number of grains per pod; medium (CV 10-20) in POG: pods with one grain, PTWG: pods with two grains; PTHG: pods with three grains, NPP: number of pods per plant, GP: number of grains per plant and YI: yield; in addition to high (CV > 20) for PFG: pods with four grains (Table 2). It was corroborated with Bohn et al. (2016), Torres, Silva and Teodoro (2015), Castro, Kouri, Alves and Silva-Neto (2014), and Ribeiro et al. (2016).

The cultivars Bônus, Fogo and ST797 showed higher averages for the variables of POG and PTWG with values of 14.90 and 24.50 units, respectively. The PTHG average was very expressive for all cultivars with the exception of Bônus, which delivered only 6.45 units plant\(^{-1}\). Bônus, Desafio, Flecha and Fogo were the genetic materials that had the highest PFG (0.93 units plant\(^{-1}\)) (Table 3). Castro et al. (2014) observed lower means of POG, PTWG and PTHG. Similar results were found by Bohn et al. (2016), which observed that soybean cultivars with higher and lower PTHG, showed statistically similar yield. Thus, these results demonstrate that the plants...
have the potential to compensate for variations in the number of pods, increasing the weight of grains, thus enabling satisfactory yield.

Table 2. Summary of analysis of variance (calculated MS and CV (%)) for POG: pods with one grain; PTWG: pods with two grains; PTHG: pods with three grains; PFG: pods with four grains; NGP: number of grains per pod; NPP: number of pods per plant; GP: number of grains per plant; and YI: yield, of soybean cultivars. Mineiros, GO, Brazil. UNIFIMES, 2020.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>POG</th>
<th>PTWG</th>
<th>PTHG</th>
<th>PFG</th>
<th>NGP</th>
<th>NPP</th>
<th>GP</th>
<th>YI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivars</td>
<td>69.48**</td>
<td>101.47**</td>
<td>81.01**</td>
<td>0.33**</td>
<td>0.10**</td>
<td>323.59**</td>
<td>1155.59**</td>
<td>819.56**</td>
</tr>
<tr>
<td>Blocks</td>
<td>0.39</td>
<td>12.72</td>
<td>13.17</td>
<td>0.00</td>
<td>0.00</td>
<td>15.23</td>
<td>96.33</td>
<td>44.47</td>
</tr>
<tr>
<td>Residue</td>
<td>2.94</td>
<td>5.50</td>
<td>10.73</td>
<td>0.02</td>
<td>0.00</td>
<td>26.94</td>
<td>161.86</td>
<td>81.09</td>
</tr>
<tr>
<td>CV (%)</td>
<td>18.79</td>
<td>12.00</td>
<td>17.57</td>
<td>3.06</td>
<td>10.83</td>
<td>11.94</td>
<td>11.68</td>
<td></td>
</tr>
</tbody>
</table>

** Significant at 1% probability by the F test.

Table 3. Averages for POG: pods with one grain; PTWG: pods with two grains; PTHG: pods with three grains; PFG: pods with four grains; NGP: number of grains per pod; NPP: number of pods per plant; GP: number of grains per plant; and YI: yield of soybean cultivars. Mineiros, GO, Brazil. UNIFIMES, 2020.

<table>
<thead>
<tr>
<th>Cultivars</th>
<th>POG</th>
<th>PTWG</th>
<th>PTHG</th>
<th>PFG</th>
<th>NGP</th>
<th>NPP</th>
<th>GP</th>
<th>YI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bônus</td>
<td>15.58 a</td>
<td>28.38 a</td>
<td>6.45 b</td>
<td>0.95 a</td>
<td>1.86 d</td>
<td>51.36 b</td>
<td>95.50 b</td>
<td>81.17 c</td>
</tr>
<tr>
<td>Desafio</td>
<td>5.11 c</td>
<td>13.70 c</td>
<td>16.86 a</td>
<td>0.85 a</td>
<td>2.36 b</td>
<td>36.53 c</td>
<td>86.51 b</td>
<td>64.88 d</td>
</tr>
<tr>
<td>Flecha</td>
<td>11.43 b</td>
<td>25.13 a</td>
<td>19.58 a</td>
<td>1.11 a</td>
<td>2.18 c</td>
<td>57.26 a</td>
<td>124.91 a</td>
<td>99.93 a</td>
</tr>
<tr>
<td>Foco</td>
<td>14.16 a</td>
<td>21.73 a</td>
<td>19.23 a</td>
<td>0.80 a</td>
<td>2.12 c</td>
<td>55.93 a</td>
<td>118.53 a</td>
<td>85.93 b</td>
</tr>
<tr>
<td>ICS7019</td>
<td>8.96 b</td>
<td>15.80 b</td>
<td>21.90 a</td>
<td>0.23 c</td>
<td>2.28 b</td>
<td>46.90 b</td>
<td>107.20 a</td>
<td>75.04 c</td>
</tr>
<tr>
<td>M5917</td>
<td>7.45 b</td>
<td>24.25 a</td>
<td>26.50 a</td>
<td>0.31 c</td>
<td>2.34 b</td>
<td>58.51 a</td>
<td>136.71 a</td>
<td>105.95 a</td>
</tr>
<tr>
<td>M7110</td>
<td>2.16 c</td>
<td>12.40 c</td>
<td>18.05 a</td>
<td>0.53 b</td>
<td>2.51 a</td>
<td>33.15 c</td>
<td>83.25 b</td>
<td>60.35 d</td>
</tr>
<tr>
<td>Power</td>
<td>3.33 c</td>
<td>12.28 c</td>
<td>16.28 a</td>
<td>0.60 b</td>
<td>2.43 a</td>
<td>32.50 c</td>
<td>79.15 b</td>
<td>55.40 d</td>
</tr>
<tr>
<td>ST721</td>
<td>8.11 b</td>
<td>18.40 b</td>
<td>21.86 a</td>
<td>0.35 c</td>
<td>2.30 b</td>
<td>48.73 b</td>
<td>111.91 a</td>
<td>75.54 c</td>
</tr>
<tr>
<td>ST797</td>
<td>14.96 a</td>
<td>23.40 a</td>
<td>19.80 a</td>
<td>0.13 c</td>
<td>2.08 c</td>
<td>58.30 a</td>
<td>121.70 a</td>
<td>66.93 d</td>
</tr>
</tbody>
</table>

Averages followed by the same letter vertically, do not differ by the Scott-Knott test, at 5% probability.

Regarding the NPP and GP variables, the cultivars Flecha, Foco, M5917 and ST797 presented the highest averages with 57.50 and 125.46 units plant$^{-1}$, in this sequence. Superior results were obtained by Torres et al. (2015) and Scheffler, Perleberg, Rodrigues and Kuhn (2016), who found NPP values above 60 pods per plant.

When the YI variable was analyzed, two cultivars Flecha and M5917 stood out, reaching a mean of 99.93 and 105.95 sc ha$^{-1}$. Similar results were found by Scheffler et al. (2016) and lower means were found by Borges et al. (2018), Bohn et al. (2016), Doná et al. (2019) attributed the averages below the climatic factor. The wide variation in yield observed between the cultivars evaluated, shows the existence of different levels of adaptability of these materials to local environmental conditions and reinforces the importance of continuing the work of evaluating cultivars (Campos et al., 2016). The agronomic performance of soybean genotypes occurs due to genetic constitution, environmental conditions and genotype x environment interaction, therefore, studies are needed to promote the proper positioning of each genotype in specific environments (Ferrari et al., 2016).

One of the genetic breeding strategies, as well as genotype positioning, is the use of Pearson’s linear correlation, in order to estimate the direction and degree of
linear association between two random traits (Olivoto et al., 2016). Pearson’s correlation coefficients arranged in the correlation network revealed 19 correlations among soybean variables. Positive correlations were diagnosed in the pairs (POGxPTWG), (POGxNPP), (POGxYI), (PTWGxNPP), (PTWGxGP), (PTWGxYI), (PTHGxNPP), (PTHGxGP), (PTHGxYI), (NPPxGP), (NPPxYI) and (NPPxYI) and negative in pairs (POGxNGP), (PTWGxNGP), (PTHGxPFG), (NGPxNPP) (Figure 1). The association between agronomic traits is important because it allows verifying the degree of interference of a trait on another of economic interest, as well as practicing indirect selection (Zuffo et al., 2016).

![Figure 1](image_url)

**Figure 1.** Network of linear correlations among traits of soybean cultivars. Significance: *5% probability and **1% probability by t test. Variables: NGP: number of grains per pod; NPP: number of pods per plant; GP: number of grains per plant; and YI: yield. Mineiros, GO, Brazil. UNIFIMES, 2019.

The use of correlation networks can increase the effectiveness of selection in soybean breeding. It allows to quickly identify the pairs of traits that present correlations of greater magnitude, to determine which groups of variables influence in a more expressive way the most important characters for the breeding program and to identify the groups of correlated variables.

The two canonical correlations and their respective canonical pairs were significant (p≤0.01) by the chi-square test (Figure 2), as well, verified in Zanatto et al. (2016). The high magnitude of the canonical correlation coefficients (r = 1.00 in the first and r = 0.91 in the second) showed a high dependency between the two groups of characters (Figure 2).

The analysis of canonical correlations showed that the increase in POG, PTWG, PTHG and NPP, in addition to the reduction in PFG and NGP, potentiate the increase in the GP variable. However, to raise the YI, it became necessary for the soybean plants to have the highest number of PTWG, PFG and NPP (Figure 2). The correlation...
studies allow to identify and quantify the associations of morphological and productive characters with the performance of the crops (Carvalho et al., 2015).

**Figure 2.** Canonical charges of primary (group 1) and secondary (group 2) yield components of soybean cultivars. Group 1: GP: number of grains per plant and YI: yield; Group 2: POG: pods with one grain; PTWG: pods with two grains; PTHG: pods with three grains; PFG: pods with four grains; NGP: number of grains per pod; and NPP: number of pods per plant. Mineiros, GO, Brazil. UNIFIMES, 2020.

From the dissimilarity matrix, it was possible to generate a dendrogram using the UPGMA clustering methodology, using the generalized Mahalanobis distance. In this methodology, the distance matrix between individuals in the population is calculated and then the most distant individuals are grouped.

Among the soybean cultivars analyzed in the dissimilarity dendrogram, two groups were generated, with a highlight for the group formed by the cultivars Desafio, M7110 and Power and the other cultivars presenting similar traits being in the second group (Figure 3). Santos et al (2015) and Rigon (2015) also found formation of different groups among soybean cultivars. The formation of groups of soybean allows, to the producers, options in the decision making for the choice of the cultivar.

The canonical axes add up to a total of explanation equivalent to 97% of the total variation of the data. Similar results were evidenced by Szareski et al. (2016), where canonical variables explained 85.05% of the existing genetic variation. The variables GP and NPP showed similarities to each other, where the cultivar Foco presented the highest GP. In the NGP variable, the closest cultivar was M7110, whereas YI was the only variable in the negative axis, where the Bônus cultivar stood out (Figure 4).

Silva et al. (2015) concluded that multivariate analysis methodologies are efficient to verify similarities or differences in yield variability, based on the chemical and physical attributes of the soil in the studied area. Also being added the influence of soybean genetic variability and seed treatment on the performance of the initial grubbing of its seedlings.

Univariate analyzes revealed significance between soybean cultivars for all variables analyzed, with Pearson's correlations ranging from 0.36 to 0.93 between positive and negative. In the multivariate, cultivar trends were diagnosed in the canonical variables tool, formation of significant groups in the canonical correlations, as well as distinctions of the dendrogram.
Figure 3. Dendrogram representative of the dissimilarity between soybean cultivars, obtained by the UPGMA clustering method, using the generalized Mahalanobis distance. The cofenetic correlation coefficient \( r \) was 0.80. Variables: NGP: number of grains per pod; NPP: number of pods per plant; GP: number of grains per plant; and YI: yield. Mineiros, GO, Brazil. UNIFIMES, 2020.

Figure 4. Analysis of canonical variables based on Mahalanobis distances from the average yield components of soybean cultivars. Variables: NGP: number of grains per pod; NPP: number of pods per plant; GP: number of grains per plant; and YI: yield. Mineiros, GO, Brazil. UNIFIMES, 2020.
CONCLUSIONS

The soybean cultivars Flecha and M5917 showed the highest yields in a tropical environment.

The cultivars showed a strong correlation between the number of grains per plant and yield.

The canonical correlations tools were efficient and complementary in the data analysis.

REFERENCES


