CORRELATIONS AND PATH ANALYSIS IN AGRONOMIC TRAITS OF SOYBEANS UNDER DEFOILIATION

CORRELAÇÕES E ANÁLISE DE TRILHA EM CARACTERES AGRONÔMICOS DE SOJA SOB DESFOLHA

Alan Mario ZUFFO¹; Francisco de ALCÂNTARA NETO²; Tiago ZOZ³; Rafael Felippe RATKE¹; Jorge González AGUILERA¹; Paulo Eduardo TEODORO¹*

1. Universidade Federal de Mato Grosso do Sul, Chapadão do Sul, MS, Brasil; 2. Universidade Federal do Piauí, Teresina, PI, Brasil; Universidade Estadual de Mato Grosso do Sul, Cassilândia, MS, Brasil. *Autor para correspondência: eduteodoro@hotmail.com

ABSTRACT: Stress events or leaf damage can alter the redistribution of photoassimilates and modify the morphological traits directly related to grain yield. This study aimed to evaluate the influence of agronomic traits on grain yield in soybean crop in function of leaf defoliation. The experiment consisted of a randomized block design with three replications, arranged in a 6 x 3 + 1 factorial scheme. The factors were composed by six reprodutive stages (R₁, R₂, R₃, R₄, R₅, and R₆) and three defoliation levels (33%, 66%, and 99%), and one additional treatment without defoliation (control). The following traits were evaluated: plant height, number of pods per plant, number of grains per pod, thousand grain weight, and grain yield. The magnitudes of the correlations between the same traits showed high oscillation according to the defoliation level. Results revealed the potential for the indirect selection for grain yield by the traits plant height and number of grains per pods under 0% of defoliation and by the trait number of pods per plant under 33% of defoliation.


INTRODUCTION

Soybean (Glycine max L. Merrill) is one of the most important oil crops worldwide, and Brazil stands out as the second largest producer in the world. In the 2016/2017 crop season, the soybean occupied an area of 33.7 million hectares, totaling a production of 103.6 million tons (CONAB, 2017). Soybean production depends on the genetic potential of cultivars and leaf damage, as they interfere with the photosynthetic intensity of the plant (MONTEIRO et al. 2017). Among the biotic factors, insect pests are one of the major problems faced by soybean producers regarding grain yield increase (SEDIYAMA, 2016).

Under field conditions, defoliating pests reduce leaf area, and therefore, the plant's photosynthetic capacity. However, the damage level depends on the time the pest remains attacking the plant, the defoliation percentage, and the phenological stage of the crop (HOFFMAN-CAMPO et al., 2012). Studies have reported that defoliation at the vegetative stages in soybean crops does not affect yield components (BAHRY et al., 2013; SOUZA et al., 2014; MONTEIRO et al., 2017). Conversely, some works have shown that defoliation at the reproductive stages affects morphological traits and grain yield (FONTOURA et al., 2006; GLIER et al., 2015; ZUFFO et al., 2015) due to the higher nutrient and photoassimilates required for grain filling.

Morphological traits directly influence soybean yield (ALCÂNTARA NETO et al., 2011), and the occurrence of stress or leaf damage in the plant can alter photoassimilates redistribution by the source-drain ratio (GODIM et al., 2008). Therefore, knowing the defoliation level that soybean plants with stand without reducing grain yield is a relevant management practice to maximize grain yield and reduce production costs. The study of this relationship is fundamental for the knowledge about assimilates partition in the plant (GODIM et al., 2008) since assimilates accumulation in the drains can be limited by the source (LOPES & LIMA, 2015).

The Pearson's correlation is a way to investigate the association between agronomic traits, allowing the quantification of the degree of linear association between two traits. Despite being widely used, the Person’s correlation is influenced by a third trait or a group of traits. To remove these undesirable effects, Wright (1921) proposed the path analysis, which considers a main dependent variable in the model and estimates the direct effects of the other variables. Therefore, it is the main analysis to study the influence of agronomic traits on grain yield. The identification of traits correlated with grain yield is an essential tool for...
Correlations and path... ZUFFO, A. M. et al.

the selection of more productive genotypes (ZUFFO et al., 2018).

Therefore, this study evaluated the influence of agronomic traits on soybean grain yield in function of plant defoliation levels, using path analysis.

MATERIAL AND METHODS

The experiment was carried out at the Center of Scientific and Technological Development in Agriculture - Fazenda Muquém / UFLA, located in the municipality of Lavras - MG (lat. 21° 4' S; long. 45° 00' W; at 918 m asl), from December 2013 to April 2014. The climate of the region is Cwa type, according to the Köppen’s classification. During the experiments, the average air temperature was 23 °C, and the rainfall was 581 mm.

The soil of the experimental area is classified as Rhodic Hapludox (Latossolo Vermelho Eutroférrico in the Brazilian classification), with the following chemical characteristics: pH in H2O of 6.4; 3.4 dag kg\(^{-1}\) of organic matter; 11.4 mg dm\(^{-3}\) of P (Mehlich\(^{3}\)); 118 mg dm\(^{-3}\) of K\(^{+}\); 5.0 cmol, dm\(^{-3}\) of Ca\(^{2+}\); 1.4 cmol, dm\(^{-3}\) of Mg\(^{2+}\); 0 cmol, dm\(^{-3}\) of Al\(^{3+}\); 2.9 cmol, dm\(^{-3}\) of H\(^{+}\) + Al\(^{3+}\); 9.6 cmol, dm\(^{-3}\) of CTC; 69.8% of base saturation and 4.7 mg dm\(^{-3}\) of S-SO\(_4^{2-}\); 0.17 mg dm\(^{-3}\) of B; 1.4 mg dm\(^{-3}\) of Cu\(^{2+}\); 34.8 mg dm\(^{-3}\) of Fe\(^{2+}\); 31.7 mg dm\(^{-3}\) of Mn\(^{2+}\); 4.9 mg dm\(^{-3}\) of Zn\(^{2+}\).

The experiment consisted of a randomized complete block design, with three replications, arranged in a 6 x 3 + 1 factorial scheme, with six reproductive stages of defoliation [(R\(_1\) (beginning of flowering), R\(_2\) (full flowering), R\(_3\) (beginning of pods formation), R\(_4\) (full pods formation), R\(_5\) (beginning of pods filling), R\(_6\) (grains 100% filled)], three defoliation levels [(33%, 66%, and 99%)], and an additional treatment without defoliation (control). Defoliation was characterized by removing one part (terminal leaflet), two parts (opposite leaflets), and three parts (all leaflets) of all the leaves developed in the plant, with the aid of scissors. Each plot consisted of four 5-m long sowing rows, spaced at 0.50 m apart, and each plot area had 10 m\(^2\) (5 m x 2 m). The two central rows were considered as useful area, excluding 1 m from each end.

Seeds were treated with pyraclostrobin + methyl thiophanate + fipronil at a dose of 2 mL kg\(^{-1}\) of seed and inoculated with Bradyrhizobium japonicum (strains SEMIA 5079 and 5080) at a dose of 3 mL kg\(^{-1}\) of seed, and sown on December 15, 2013. Fertilization consisted of 350 kg ha\(^{-1}\) of the N-P\(_2\)O\(_5\)-K\(_2\)O (02-30-20) formulation, which was applied to the groove. The soybean cultivar BRS Favorita RR\® was used in the experiment at a seed density of 12 plants per meter of row, which received the established defoliation levels.

During the plant development, weeds, pests, and diseases management were performed when necessary, by applying Glyphosate, Pyraclostrobin + Epoxiconazole, Azoxystrobin + Cyproconazol, Teflubenzuron, Chlorpyrifos, and Cypermethrin.

At the harvest time, plant height was determined from the soil surface to the apex of the apical meristem, using a millimeter ruler. Subsequently, ten plants per plot were collected to evaluate the number of pods per plant and number of grains per pod. The plants contained in the useful area of the plot were manually collected and mechanically processed, and the production was transformed into kg ha\(^{-1}\) of grains, standardized at 13% moisture. The variable thousand seed weight was determined according to Brasil (2009).

Initially, the analysis of variance was performed for the studied variables, and afterward, the Pearson’s correlation (r) was carried out between the pairs of traits, based on equation 1:

\[
r = \frac{COV_{XY}}{\sqrt{\sigma_X^2 \times \sigma_Y^2}}
\]

Where \(COV_{XY}\) is the covariance between traits X and Y; \(\sigma_X^2\) is the variance of trait X; and \(\sigma_Y^2\) is the variance of trait Y. A correlation network was used to graphically express the functional relationship between estimates of the phenotypic correlation coefficients between traits, where the proximity between nodes (traces) was proportional to the absolute value of the correlation between those nodes. The thickness of the edges was controlled by applying a cut-off value of 0.60, which meant that only if \(|r_{ij}| \geq 0.60\) had their edges highlighted. Finally, positive correlations were highlighted in blue, and negative correlations were represented in red.

Subsequently, a multicollinearity diagnosis of the X’X correlation matrix was performed. When the multicollinearity was moderate (condition number< 100), based on the classification of Montgomery & Peck (2001), a constant \(k = 0.10\) was added to the diagonal of the X’X matrix to perform the path analysis, resulting in weak multicollinearity (condition number<100).

Path analysis, which considered grain yield (GY) as the main dependent variable, was performed using the model described in equation 2:

\[
GY = \beta_1 PH + \beta_2 NGP + \beta_3 TSW + \beta_4 NFP + p_0
\]

(Eq.2)
Correlations and path…

Where $\beta_1, \beta_2, \ldots, \beta_{11}$ were the direct effect estimates for traits PH, NGP, TSW, and NPP on GY; and $\hat{\rho}_e$ was the residual effect of the analysis. Thus, a normal system of equations was used to estimate the direct and indirect effects of each explanatory variable on GY, as given by equation 3:

$$
\begin{bmatrix}
1 & \vdots & \vdots & \vdots & 1 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
1 & \vdots & \vdots & \vdots & 1 \\
\end{bmatrix}
\begin{bmatrix}
\beta_1 \\
\vdots \\
\vdots \\
\vdots \\
\beta_{11} \\
\end{bmatrix}
= 
\begin{bmatrix}
\epsilon_1 \\
\vdots \\
\vdots \\
\vdots \\
\epsilon_{11} \\
\end{bmatrix}
$$

(Eq.3)

The determination coefficient ($R^2$) from the path analysis was obtained by equation 4:

$$
R^2 = \beta_1 \beta_{DP.GY} + \cdots + \beta_n \beta_{NPP.GY}
$$

(Eq.4)

The residual effect ($\hat{\rho}_e$) of the path analysis was obtained by equation 5:

$$
\hat{\rho}_e = \sqrt{1 - R^2}
$$

(Eq.5)

All the analyses were performed in the statistical software Genes version 5.1 for Windows (Software of Statistical Analysis, GENES, Viçosa, MG, BRA).

RESULTS AND DISCUSSION

Results revealed a significant effect of defoliation on the number of pods per plant (NPP), number of grains per pod (NGP), thousand seed weight (TSW), and grain yield (GY) (Table 1). The influence of the defoliation percentage on soybean production components was also verified by Glier et al. (2015) and Monteiro et al. (2017). The stage effect was significant only for NPP and GY. The defoliation x stage interaction revealed a significant effect only for grain yield. The Factorial vs. Additional contrast was significant for all the evaluated traits. These results indicate that all traits presented differences between the treatments in relation to the control. This difference occurred mainly due to the significant effect of the defoliation factor. Therefore, correlations and path analysis were performed for each defoliation level.

Table 1. Analysis of variance of the data for plant height (PH), number of pods per plant (NPP), number of grains per pod (NGP), thousand seed weight (TSW), and grain yield (GY) obtained in the trials of defoliation levels and reproductive stages of defoliation in a soybean crop of cultivar BRS Favorita RR. Lavras, MG, Brazil. 2013/14 crop season.

<table>
<thead>
<tr>
<th>Sources of variation</th>
<th>PH (cm)</th>
<th>NPP (unit)</th>
<th>NGP (g)</th>
<th>TSW (Kg ha⁻¹)</th>
<th>GY (Kg ha⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks</td>
<td>4.73</td>
<td>6.50</td>
<td>0.16</td>
<td>46.02</td>
<td>91048.09</td>
</tr>
<tr>
<td>Defoliation (D)</td>
<td>57.45**</td>
<td>352.28**</td>
<td>0.76**</td>
<td>1827.80**</td>
<td>6012007.84**</td>
</tr>
<tr>
<td>Stage (S)</td>
<td>30.14**</td>
<td>74.09**</td>
<td>0.14ns</td>
<td>186.49ms</td>
<td>2538777.79**</td>
</tr>
<tr>
<td>D x S</td>
<td>18.50ns</td>
<td>16.67ns</td>
<td>0.17ns</td>
<td>109.22ms</td>
<td>104324.64*</td>
</tr>
<tr>
<td>Factorial vs. Additional</td>
<td>116.28**</td>
<td>591.75**</td>
<td>591.75*</td>
<td>599.21*</td>
<td>1387275.47**</td>
</tr>
<tr>
<td>Treatments</td>
<td>31.97ns</td>
<td>138.19**</td>
<td>0.23ns</td>
<td>347.39**</td>
<td>870450.96**</td>
</tr>
<tr>
<td>Residue</td>
<td>24.06</td>
<td>10.61</td>
<td>0.11</td>
<td>125.42</td>
<td>45349.94</td>
</tr>
<tr>
<td>CV (%)</td>
<td>5.37</td>
<td>7.77</td>
<td>18.56</td>
<td>8.06</td>
<td>11.80</td>
</tr>
</tbody>
</table>

** and * significant at 1 and 5% probability level by the F test, respectively. ** - not significant; DF – degrees of freedom; CV – coefficient of variation.

Figure 1 shows the correlation network between the traits for the different defoliation levels, evidencing considerable changes in the magnitude and direction of the correlation coefficients. Under 0% of defoliation (Figure 1a), GY was positively correlated with PH and NGP. These results are partially similar to those of Zuffo et al. (2018), who verified a positive and significant association between grain yield and plant height in soybean plants. However, the correlation was null for the other defoliation levels. The fact that the magnitude and direction of the correlation coefficients change according to the increase in the defoliation percentage can be related to the change in the photoassimilate redistribution pattern from the source regions to the drains. For Alexandrino et al. (2005), defoliation is one of the factors that interfere with assimilates partitioning, whose effect on plant growth depends on the intensity of the leaf area removal and the ability of the plant to reestablish new assimilates partition from the new sources to the new drains. Moreover, after defoliation, the photosynthates recently synthesized in the residual leaves and the organic reserves participate in the leaf area reconstruction, prioritizing the terminal meristems (SCHNYDER; DE VISSER, 1999).
TSW was negatively correlated with GY, PH, and NGP under 0% defoliation (Figure 1a). Nevertheless, the TSW x NGP correlation was positive and of high magnitude under 66% of defoliation (Figure 1c). Similarly, the GY x NPP correlation at 0% of defoliation (Figure 1a) was negative; however, at 33% of defoliation (Figure 1b), the correlation was positive and of high magnitude. These results confirm the hypothesis of a change in the relationship between traits in function of the defoliation level in soybean, possibly due to photoassimilates partition.

Thus, despite its importance, the Pearson's correlation coefficient can lead to misunderstandings about the relationship between two traits and may not be a true measure of cause and effect. A high or low correlation coefficient between two traits may be the result of the effect of a third trait or a group of traits on the pair, not giving the exact relative importance of the direct and indirect effects of these factors (CRUZ et al., 2012). Therefore, path analysis was carried out to investigate the cause and effect relationship. TEODORO et al. (2014) report that this analysis provides detailed knowledge of the influences of the traits involved and justify the existence of positive and negative correlations, of high and low magnitude, between the studied traits. Consequently, the traits that can be used for indirect selection at each defoliation level must be identified.

For the path analysis under 0% of defoliation, multicollinearity was corrected by adding the constant $k = 0.10$ to the diagonal of the $X'X$ matrix. In the presence of multicollinearity, the variances associated with the path coefficient estimators can reach exceedingly high values, becoming unreliable. Moreover, parameters estimates may assume absurd values or values inconsistent with the biological phenomenon (CRUZ et al., 2012).

For 0% of defoliation (Figure 2), traits that presented the highest direct positive effect on GY were PH and NGP; conversely, TSW had a negative effect. These traits can be used in the indirect selection of more productive soybean genotypes under 0% of defoliation since the direction of direct effects was the same when compared with their correlations with GY (Figure 1).
Figure 2. Direct and indirect effect of four traits on soybean grain yield. Lavras, MG, Brazil. 2013/2014 crop season. 0% of defoliation. Determination coefficient of causal model ($R^2$): 0.967.

Effect of the residual variable (ERV): 0.180. Multicollinearity test of explanatory variables: Weak (NC<100), $k$ value: 0.100.

For 33% of defoliation (Figure 3), NPP had the highest direct effect on GY. NPP was also the most promising for indirect selection for presenting positive correlation in the same direction of its direct effect on GY. These results may be useful in soybean breeding programs that aim at pest-tolerant genotypes since the selection of genotypes with higher NPP will result in higher GY (Figure 1).

Figure 3. Direct and indirect effect of four traits on soybean grain yield. Lavras, MG, Brazil. 2013/2014 crop season. 33% of defoliation. Determination coefficient of causal model ($R^2$): 0.821. Effect of the residual variable (ERV): 0.422. Multicollinearity test of explanatory variables: Weak (NC<100), $K$ value: 0.100.

PH can be used at 66 and 99% of defoliation (Figure 4 and 5, respectively). However, in both cases, the variables with the highest direct effects on GY did not present low magnitude correlations with GY (Figure 1). Therefore, other variables should be evaluated for indirect selection of soybean genotypes.
Correlations and path…

ZUFFO, A. M. et al.

...genotypes that are more productive under these conditions. The determination coefficient obtained in all cases was higher than 0.70, indicating that more than 70% of the GY variation, at all defoliation levels, was explained by the variables evaluated.

![Figure 4](image)

**Figure 4.** Direct and indirect effect of four traits on soybean grain yield. Lavras, MG, Brazil. 2013/2014 crop season. 66% of defoliation. Determination coefficient of causal model ($R^2$): 0.700. Effect of the residual variable (ERV): 0.546. Multicollinearity test of explanatory variables: Weak (NC<100). $K$ value: 0.100

![Figure 5](image)

**Figure 5.** Direct and indirect effect of four traits on soybean grain yield. Lavras, MG, Brazil. 2013/2014 crop season. 99% of defoliation. Determination coefficient of causal model ($R^2$): 0.855. Effect of the residual variable (ERV): 0.379. Multicollinearity test of explanatory variables: Weak (NC<100). $K$ value: 0.100
Zuffo et al. (2017) reported direct effects on soybean grain yield. The authors cite plant height at harvest, insertion of the first pod, number of pods, and number of grains per pod as the responsible for these effects. Conversely, Alcantara Neto et al. (2011) verified that the number of pods per plant is the production component with the highest direct effect on dry matter and total grain yield per plant in soybean crops. Zuffo et al. (2018) observed that number of pods per plant and number of grains per pod had high direct effects on soybean grain yield. These studies did not evaluate defoliation in soybean crops, and therefore, such divergences may be related to the intrinsic characteristics of the cultivars used.

Results of correlations by the defoliation level indicate the variability among the soybean crop and demonstrate that the performance of the soybean crop was not coincident at the different defoliation levels. Thus, breeding programs intended to select pest-tolerant genotypes should perform path analysis by the defoliation level. Although soybean crops have a high capacity to adapt to different environmental and management conditions by changing plants morphology, architecture, and yield components (Pires et al., 2000), defoliation levels change the magnitudes of correlations between the same traits, leading to high oscillation.

CONCLUSIONS

The magnitudes of correlations between the same traits showed high oscillation according to the defoliation level.

Indirect selection for grain yield will not be effective at all defoliation levels.

The potential of indirect selection was detected for grain yield by the traits plant height and number of grains per pod at 0% of defoliation and by the trait number of pods per plant at 33% of defoliation.

The results for the defoliation levels of 66% and 99% indicated that other variables should be evaluated for indirect selection of more productive soybean genotypes under these conditions.

RESUMO: Ocorrências de estresse ou danos foliares podem levar a alteração na redistribuição de fotoassimilados e modificar os caracteres morfológicos envolvidos diretamente na produtividade dos grãos. Portanto, objetivou-se avaliar por meio da análise de trilha, a influência de caracteres agronômicos na produtividade de grãos na cultura da soja em função de níveis desfolha da planta. O delineamento experimental utilizado foi o de blocos casualizados com três repetições, dispostos em esquema fatorial 6 x 3 + 1. O primeiro fator foi composto por seis estádios reprodutivos de desfolha (R₁, R₂, R₃, R₄, R₅, R₆) x três níveis de desfolha (33%, 66%, 99%) e um tratamento adicional sem desfolha (controle). Os caracteres avaliados foram: altura de plantas, número de legumes por planta, número de grãos por legume, peso de mil grãos e produtividade de grãos. As magnitudes das correlações entre os mesmos caracteres apresentaram alta oscilação de acordo com o nível de desfolha, sendo que, existe potencial da seleção indireta para produtividade de grãos por meio da altura de plantas e número de grãos por legumes sob 0% desfolha e do número de legumes por planta sob 33% de desfolha.


REFERENCES


Correlations and path…


Correlations and path…

ZUFFO, A. M. et al.


