AgroSym
2022

BOOK OF PROCEEDINGS

XIII International Scientific Agriculture Symposium
"AGROSYM 2022"
October 6-9, 2022

agro **2022**
sym

MULTIVARIATE ANALYSIS OF AGRONOMIC TRAITS IN MID-SEASON SOYBEAN VARIETIES

Vesna PERIĆ*, Mirjana SREBRIĆ, Danijela RISTIĆ, Marijenka TABAKOVIĆ, Valentina NIKOLIĆ, Snežana MLADENOVIĆ DRINIĆ

*Maize Research Institute „ZemunPolje", SlobodanaBajica 1, Belgrade, Serbia
*Corresponding author: vperic@mrizp.rs

Abstract

Principal Component Analysis (PCA) is a useful tool for processing multiple data, which are often encountered in breeding practice. This method is suitable for the evaluation of genotypes on the basis of multiple traits and graphical presentation of relationships between traits. This study included 16 soybean mid-season genotypes (maturity group I), originated from different regions of the world, maintained in soybean collection of Maize Research Institute Zemun Polje. Field trials were carried out at two locations, during two years, according to a RCB design with three replications. The genotypes were evaluated in respect to eight major agronomic traits: *PH* – plant height, *NN* – node number, *PN* – pod number, *SN* – seed number, *TSW* – 1000 seed weight, *SYP* – seed yield per plant, *PROT* – protein content, *OIL* – oil content. First two PCA axes encompassed a large portion of the variance of standardized data (75,9%). Biplot distinguished genotypes of potential importance for various breeding targets. Two genotypes stood out with the largest 1000 seed weight. One variety formed a larger number of pods and the seed number per plant as compared to the group average, achieving the highest grain yield per plant. Two genotypes were among the most productive ones, with a larger number of pods as well as a higher 1000 seed weight, compared to the average. The most promising variety was Laura, which had a high yield and higher protein content than the average, and could be used as a potential germplasm source for the simultaneous improvement of both traits. Correlations among traits determined by PC biplot were in accordance with Pearson's correlation coefficients.

Keywords: *soybean, quantitative traits, multivariate analysis, correlations*

Introduction

The areas under soybean in Serbia are continuously increasing in the last decade. The total soybean harvested area in Serbia in 2020 was 236758 ha, with an average yield 3,175 t/ha (FAOSTAT, 2022). Although the seed yield is one of the most important traits in soybean breeding programs, released cultivars have to adjust to the requirements of processing industry regarding chemical composition and technological quality of the grain (Peric *et al.,* 2018).The success of breeding programs depends on the availability of a genetic pool with adequate diversity (Gwinner *et al.,* 2017), and possibility to identify potential sources for the enhancement of agronomic traits through different breeding methods. Promising genotypes should simultaneously unite a set of positive traits, which could elevate yield to fulfil the market demands (Cruz, 2013). Multivariate exploratory techniques, such as PCA (Principal Component Analysis) can be used for selecting superior genotypes, simultaneously analyzing important agronomic traits and the relationships among them, as well as identifying the most influencing traits in a selection process (Leite *et al.,* 2018). PCA is a widely used method for evaluation of

genotypes regarding multiple traits in soybean and other crops (Yan and Rajcan, 2002; Mohamadi and Amri, 2011; Mirosavljevic *et al*., 2015; Peric *et al*., 2018).

The identification of genotypes that are at the same time superior for negatively correlated traits presents a special challenge. The aim of this study was: to evaluate genotypes on the basis of multiple traits and identify potential breeding sources among them; examine the correlations between agronomically important traits and suggest a reliable selection criterion by applying the method of PCA.

Material and methods

Experimental material for this study included sixteen soybean accessions belonging to maturity group I (mid-season genotypes), maintained in soybean collection of Maize Research Institute "Zemun Polje". Part of the examined genotypes was developed in Serbia (domestic ones), while the other part originated from different geographical regions of the world (introductions). The field trials were set up during two growing seasons (2011 and 2012), at two locations in Serbia (Zemun Polje and Pančevo), according to a randomized complete block design with 3 replications. The experimental plot size was 5 m^2 with two rows per plot The standard agricultural practice was applied. The soil type at the Zemun Polje location was slightly calcareous chernozem, i.e. carbonate chernozem on the loess terrace at the Pančevo locality. During the both research years, total rainfalls were significantly lower and average temperatures much higher than multiyear average, so 2011 was characterized as moderately dry, while 2012 was marked as extremely dry year. The samples consisted of 30 plants per genotype were collected at the R8 stage, by random selection of 30 plants per genotype, and scored for eight agronomically important traits: *PH* –plant height, *NN* – node number, *PN* – pod number, *SN* – seed number, *TSW* – 1000 seed weight, *SYP* – seed yield per plant, *PROT* – protein content, *OIL* – oil content. Seed samples were analyzed for protein and oil content (% on a dry matter basis) with grain analyzer Infraneo, Chopin Technologies®, based on the near-infrared transmission spectroscopy (*NIRT*). Principal Component Analysis (PCA) was performed as the data reduction technique, in order to identify minimum number of traits which contributes to maximum variation, and to rank genotypes on the basis of PC scores. Pearson's correlation coefficients between agronomically important traits were calculated and compared to correlations displayed by PC biplot.

Results and Discussion

Two axes of PCA biplot (PCA1 and PCA2) of soybean genotypes maturity group I explained 75.9% of the variance of the standardized data (Figure 1). Dispersion of genotypes along the first axis (PC1) was mainly based on *PROT*, *OIL*, *PH*, *NN*, *PN* and *SN*, while second axis (PC2) separated genotypes by *TSW* and to the lesser extent by *SYP*.

Genotypes intermediate for most of the analysed traits (Hodgson 78 and Danubian) had short vectors and were positioned near the coordinate origin. Biplot highlighted two cultivars with the largest *TSW*- Balkan and Ardin. The Balkan genotype formed a larger *PN* and *SN* per plant, as compared to the group mean, achieving the highest *SYP*. The high yield of this genotype was expected, considering that it was a domestic variety used as a standard in trials for new varieties registration in maturity group I, until 2013. Genotypes Brock and Parker ranked among the most productive genotypes, forming higher *PN* as well as *TSW* compared to the average, while for the

SN per plant these two varieties were mostly intermediate. Cultivar Laura also stood out in terms of high yield and *TSW* as compared to the group average, while it was intermediate for the *PN* and *SN* per plant. High-yielding genotypes were mostly intermediate for protein content and synthesized high oil content, except for the Laura variety, which had a higher yield and higher than average protein content, representing a possible source of germplasm for simultaneous improvement of both traits. Genotypes Ika, Shine, A 1937, Ravnica and NK 15 50 formed a very high *SN* and *PN* per plant, but by *TSW* these genotypes were below the group average, while for *SYP* mostly intermediate or slightly above the group average, as well as for the *OIL*.

Varieties Chornaya, Daniela and Krizia had *PROT* significantly above the group average, but the lowest grain yield and the lowest *OIL* made these genotypes less important for breeding for simultaneous improvement of several traits. Nevertheless, these 3 genotypes can be used in breeding for improved seed protein content.
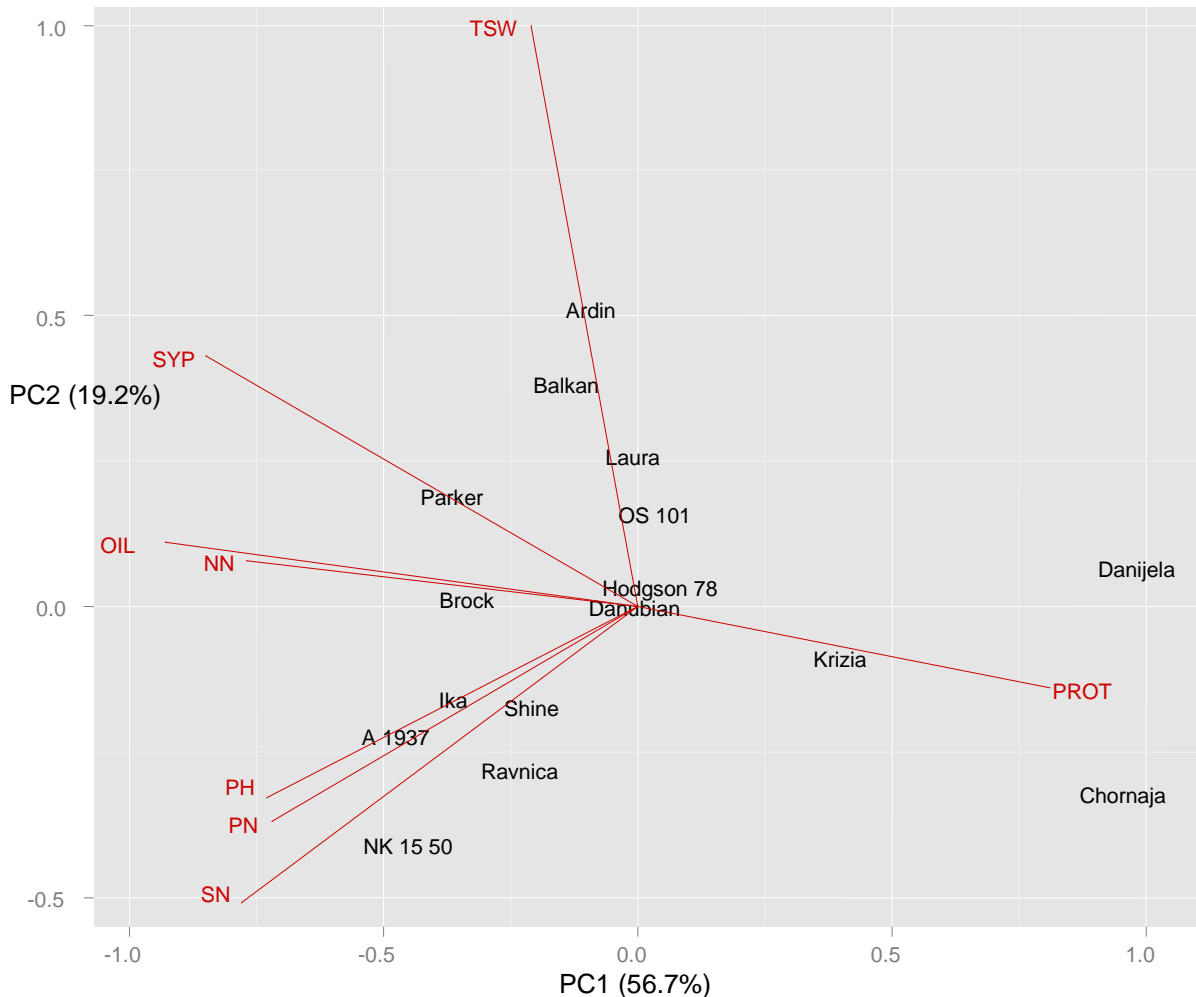

Figure 1. PCA biplot of 16 soybean genotypes maturity group I evaluated for important agronomic traits

The angles between trait vectors on PCA biplot figured out that the *SYP* showed positive correlation with *OIL*, *NN*, *PH*, and *PN* and *SN* per plant while negatively associated with *TSW* and *PROT*. The correlations between the traits shown by the biplot were in accordance with *Pearson's* correlation coefficients (Table 1), regarding vector directions (correlation sign) and the angles between vectors (correlation strength), as well.

The highest correlation coefficient was determined between *OIL* and *PROT*, indicating negative and highly significant association (-0.91***). *GYP* showed negative high significant correlation with *PROT* (-0.69**) and positive highly significant correlation with *OIL* (0.75**). Negative correlation of yield and protein, on one hand, and positive between yield and oil, on the other, as well as negative association of protein and oil has been confirmed in many studies (Taški-Ajdukovic *et al.,* 2010; Li and Burton, 2002, Popovic *et al.,* 2013), indicating a significant problem in development of high yielding soybean varieties with a satisfying level of protein and oil in grain. Similar findings were reported by Peric *et al.* (2018), *GYP* was in a highly positive and highly significant correlation with *TSW* (0.63**), and a medium positive significant correlation with the *SN* (0.56*) and *PN* per plant (0.55*). By comparing the correlation coefficients between yield and three main yield components (*SN*, *PN* and *TSW*), it could be concluded that *TSW* was the component with the greatest influence on yield. At the same time, the very weak negative and insignificant association of *TSW* with *PN* (-0.34) and *SN* (-0.28), suggests that selection of larger seed genotypes within maturity group I could possibly improve yield without significantly reducing the number of seeds and number of pods per plant.

Table 1. *Pearson's* correlation coefficients between the traits

	NN	PN	SN	TSW	SYP	PROT	OIL
PH	0.72**	-0.23	-0.06	0.20	0.04	-0.06	-0.05
NN		-0.06	0.13	0.29	0.25	0.16	-0.11
PN			0.81**	-0.32	0.51**	-0.21	0.17
SN				-0.15	0.78**	-0.43*	0.07
TSW					0.49**	0.08	-0.29
SYP						-0.37*	-0.09
PROT							-0.57**

PH –plant height, *NN* – node number *PN* – pod number, *SN* – seed number, *TSW* – 1000 seed weight, *SYP* – seed yield per plant, *PROT* – protein content, *OIL* – oil content
*P < 0.05; ** P <0.01

Conclusion

Application of PCA for simultaneous analysis of important agronomic traits in mid-season soybean genotypes revealed potential valuable sources for breeding for different goals. Five higly-productive genotypes were identified, superior by both yield and yield components. The most promising variety was Laura, which had a high yield and higher protein content than the average, and has been used in other breeding programs, presenting a potential germplasm source for the simultaneous improvement of both traits. Correlations among agronomic traits based on PC biplot were in agreement with *Pearson's* correlation coefficients, suggesting that *TSW* could be efficient selection criterion in breeding for soybean seed yield.

Acknowledgements

References

Cruz C.D. (2013): Genes: a software package for analysis in experimental statistics and quantitative genetics. ActaScientiarum. Agronomy, 35 (3), 271-276.

FAOSTAT (2022): Food and Agriculture Organization (available at https://www.fao.org/faostat/en/#data/QCL

Gwinner R., Setotaw T.A., Pasqual M., Santos J.B., Zuffo A.M., Zambiazzi E.V., Bruzi A.T., (2017): Genetic diversity in Brazilian soybean germplasm. Crop Breeding and Applied Biotechnology, Viçosa, MG, 17(4), 373-381.

Leite W.S., Unêda-Trevisoli S.H., da Silva F.M., da Silva A. J., Di Mauro A.O. (2018): Identification of superior genotypes and traits of soybean through of multivariate analysis and selection index. Revista Ciência Agronômica, 49 (3), 491-500.

Li H., J.W. Burton (2002):Selecting increased seed density to increase indirectly soybean seed protein concentration Crop Sci. 42, 393 – 398.

Mirosavljevic M., Przulj N., Canak P., Momcilovic V., Acin V., Jockovic B., Hristov M., Mladenov N. (2015): Field and Vegetable Crops Research 52(2), 74-79.

Mohamadi, R., A. Amri (2011): Graphyc analysis of trait relations and genotype evaluation in durum wheat. Journal of Crop Improvement 25, 680–696.

Perić V., Srebrić M., Nikolić A., Mladenović-Drinić S. (2018): Application of Multivariate Analysis for Genotype Evaluation in Soybean. Breeding Grasses and Protein Crops in the Era of Genomics, Springer Int. Pub. Ag, 2018, 215-219.

Popović V., Miladinović J., Vidić M., Tatić M., Sikora V., Ikanović J., Dozet G. (2013): Productive and Quality Characteristics of Soybean in Agroecological Conditions of Sombor,Serbia.Field and Vegetable Crops Research 50(2), 67-74.

Taski-Ajdukovic K., Djordjevic V., Vidic M., Vujakovic M. (2010): Subunitcomposition of seed storage proteins in high protein soybean genotypes. Pesq.Agropec. Bras., Brasilia, 45 (7), 721-729.

Yan, W., I. Rajcan, (2002): Biplot analysis of test sites and trait relations of soybean in Ontario. Crop Science 42, 11-20.