EUROSOYA, 1996, 26-33 ISSN 0760-0224

Path coefficient analysis of the effect of yield, oil content and the duration of vegetative and reproductive periods on seed protein content in Soybean

Miladinovic J.*, Hrustic M., Vidic M. and Tatic M.

Research Institute of Field and Vegetable Crops, M. Gorkog 30, 21000 Novi Sad, Yugoslavia

* corresponding author

Abstract

During 1991 and 1992, a study was carried out to determine the effect of yield, oil content, and the duration of the vegetative and reproductive periods on protein content in soybean seeds. In 1991 protein content was highly negatively correlated with yield, oil content, and the duration of the vegetative period; in 1992 only the duration of the vegetative period was associated with seed protein. Path-coefficient analysis showed that in 1991 only oil content had a significant direct effect on protein content, whereas in 1992 only length of the vegetative period had a direct effect on seed protein content. Our findings have shown that, despite its high heritability, protein content primarily depends on environmental factors and other traits, such as oil content.

Keywords: Soybean, Protein content, Path coefficient analysis

INTRODUCTION

The value of grain-type soybeans (*Glycine max* (L.) Merr.) lies in their oil and protein content. Mature seeds are usually 40% protein, 20% oil, 17% cellulose and hemicellulose, 7% sugars, 5% crude fiber, and about 6% ash on a dry-weight basis (Rubel et al., 1972). Soybeans are the most important source of plant protein, which is, in turn, an essential constituent of the animal diets. Having this in mind, it is only natural that increasing grain protein content should be a major objective of soybean breeding.

Past researches have shown that protein content is primarily a result of additive gene action (Brim and Cockerham, 1961; Hanson et al., 1967; Ishige, 1984; McKendry et al., 1985), although data pointing to non-additive components of variability for protein percentage are not rare, either (McKendry et al., 1985; Sabbouh and Edwards, 1986). Generally speaking, protein content is more heritable than any other quantitative trait, its heritability values ranging between 0.51 and 0.96 (Brim, 1973; Shannon et al., 1972; Shorter et al., 1976; Burton and Brim, 1981; Openshaw and Hadley, 1984).

In spite of its high heritability, the major difficulty in breeding programs for increased protein content of soybean seed has been the usually negative genetic correlations between percentage seed protein and seed yield (Thorne and Fehr, 1970; Hartwig and Hinson, 1972; Shannon et al., 1972; Shorter et al., 1976; Simpson and Wilcox, 1983; Burton, 1987; Leffel, 1988; Holbrook et al, 1989) as well as the high negative genetic correlations between seed protein and other economically important traits, such as seed oil (Hymowitz et al., 1972; Taira and Taira, 1976; Leffel, 1988). Therefore, progress in increasing protein content of soybean seed using various methods of selection (Thorne and Fehr, 1970; Erickson et al, 1981; Openshaw and Hadley, 1984; Sebern and Lambert, 1984; Wehrmann et al., 1987) has been slow, meaning that this trait, just like any other quantitative trait, is dependant on a number of different factors.

This study represents an attempt to obtain a clearer picture of the effect of some factors on seed protein content in soybean.

MATERIAL AND METHODS

The study was conducted during growing season of 1991 and 1992 at the Experiment Field of the Institute of Field and Vegetable Crops at Rimski Šanèevi. It included 12 promising soybean lines, nine of which have now become registered varieties and have come to cover the largest portion of the acreage in soybean in our country.

A randomized complete block design with five replications was used and the basic plot size was 10 m^2 (four rows five meters long, with 0.5 m between rows). The sowing was done on April 23, 1991 and April 24, 1992 and the harvesting at full maturity. In order to avoid border row effect, only two middle rows, 3m in length were harvested manually from each basic plot.

Seed yield was expressed as kg/ha and adjusted to 14% moisture. Oil and protein content was measured by means of an INFRATEK-1225, a device based on infrared spectroscopy, and then converted to dry matter (0 % moisture). The length of the vegetative period was determined (the number of days from emergence to the appearance of the first flower, VE - R1), as well as that of the reproductive period (the number of days from the appearance of the first flower to full maturity, R1 - R8) (Fehr and Caviness, 1977). In order to determine differences among genotypes in each of the year for each of the traits measured, obtained data were statistically processed using ANOVA (Singh and Chaundry, 1979).

Results were statistically processed using Path-coefficient analysis, a method which enables the study of direct and indirect effects of the independent variables (x_1, x_2,...x_k) on the dependent variable (y), as well as the study of the proportion of their joint action (Singh and Chaundry, 1979). The subject of this study was the effect of yield (x_1), oil content (x_2), and the duration of the vegetative (x_3) and the reproductive (x_4) period on seed protein content in soybean (y). All calculations were made using GEN 1.0., a computer program for quantitative genetic analysis.

Trial was set only about 500 m away from meteorological station at Rimski Šanèevi, so the data regarding the temperatures and rainfall during the 1991 and 1992 growing seasons were obtained from there monthly (figure 1).

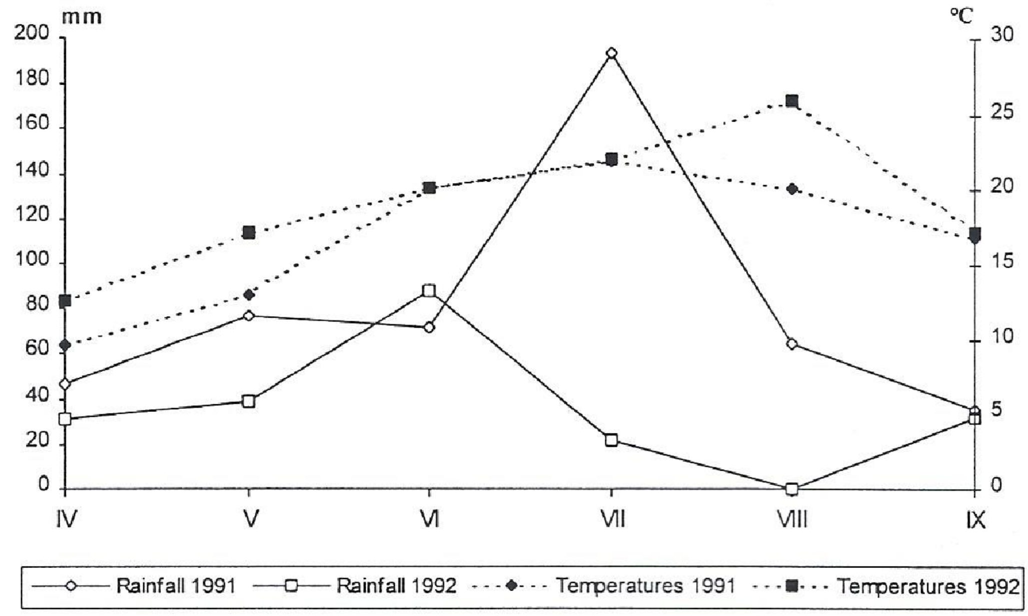

Figure 1:- Monthly mean air temperatures (°C) and monthly total rainfall (mm) in vegetation season in 1991 and 1992

RESULTS AND DISCUSSION

This experiment was replicated in each of two years. A statistical analysis of the data is essential to determine if differences among genotypes were significant in each of the years for each of the traits measured.

Table 1: Mean values of investigated traits in 1991 and 1992

Genotype	MG	Yield (kg/ha)		Protein content (% DM)		Oil content (% DM)		Vegetative period duration (days)		Reproductive period duration (days)	
		1991	1992	1991	1992	1991	1992	1991	1992	1991	1992
1. Jelica	00	3456*	2472	42.3**	37.8*	15.1**	19.3**	44**	37	72**	70**
2. NS - L - 82	00	3216**	2550	43.9**	39.3	13.8**	19.1**	42**	37	80**	70**
3. Afrodita	0	3938	2768*	39.4	38.4	18.5**	20.1**	48**	37	77	70**
4. Balkan	I	4058	2356	40.4	38.3	18.6**	17.3**	50*	38	78	79**
5. NS-L-1006	I	3688	2040	41.0	39.7	16.1**	16.2**	46**	37	82**	96**
6. NS-L-1007	I	3538*	2138	40.2	38.1*	15.7**	17.7	47**	37	82**	83
7. Ravnica	**I**	**4118**	**2274**	**39.3**	**39.4**	**17.6**	**18.0**	**52**	**37**	**77**	**83**
8. NS-L-2026	I	4262	2124	40.8	39.5	16.8*	18.2	52	37	78	83
9, Šumadija	II	3772	2514	38.9	36.4**	18.7**	19.1**	49**	39**	87**	96**
10. Simonida	II	3332*	2246	40.6	36.5**	17.4	18.8**	49**	41**	85**	80**
11. Nizija	II	4042	2232	38.7	37.2**	19.1**	19.1**	49**	40**	81**	82
12. Vojvoðanka	II	4132	2324	39.1	37.1**	19.2**	18.4	51	40**	83**	81*
LSD 0,05		510.9	438.0	2.02	1.18	0.67	0.48	1.76	1.64	1.44	1.76
0,01		694.3	595.3	2.74	1.60	0.91	0.65	2.39	2.23	1.96	2.39

Variety Ravnica was taken as a basis of comparison, because m.g. I is optimal for environmental conditions of Yugoslavia, and this variety occupies largest portion of total soybean acreages in our country.

Analysis of Variance showed that there were significant differences among genotypes in each of the years for each of the traits measured, and therefore, further analysis make sense (Table 1).
All the genotypes under investigation had a higher yield and a higher protein content in 1991 than in 1992. The highest yield was recorded in the line NS-L-2026 in 1991 (4.262 kg/ha) and the lowest in the line NS-L-1006 in 1992 (2.040 kg/ha). The highest protein content was found in the line NS-L-82 in 1991 (43.9%) and the lowest in Šumadija in 1992 (36.4%). In all of the genotypes, the vegetative period was longer in 1991 than in 1992; reproductive period were longer for about half the genotypes in 1992. Oil percentages were generally higher in 1992, but the total oil yield per ha was greater in 1991 because of higher seed yields in 1995. The highest oil content was found in Afrodita in 1992 (20%) and the lowest in Jelica in 1991 (15.1%).

Based on these results, the interdependence of the traits under investigation was studied using Path coefficient analysis.

Table 2 : Genetic correlations between investigated traits in 1991 (above diagonal) and in 1992 (below diagonal)

Traits	1	2	3	4	y
1. Yield	1.000	$r_{12} = 0.706^{**}$	$r_{13} = 0.829^{**}$	$r_{14} = -0.135$	$r_{1y} = -0.672^{**}$
2. Oil content	$r_{12} = 0.774^{**}$	1.000	$r_{23} = 0.761^{**}$	$r_{24} = 0.307$	$r_{2y} = -0.899^{**}$
3. Vegetative period	$r_{13} = -0.098$	$r_{23} = 0.200$	1.000	$r_{34} = 0.171$	$r_{3y} = -0.761^{**}$
4. Reproductive period	$r_{14} = -0.571^{*}$	$r_{24} = -0.589^{*}$	$r_{34} = 0.180$	1.000	$r_{4y} = -0.352$
y. Protein content	$r_{1y} = -0.207$	$r_{2y} = -0.440$	$r_{3y} = -0.817^{**}$	$r_{4y} = -0.081$	1.000

Significantly at level 0.05 (*) and 0.01 (**)

Table 2 shows the coefficients of correlation between the dependent variable and the independent variables (r_{y1}, r_{y2},...r_{yk}) as well as between the independent variables themselves (r_{12}, r_{13},...$r_{k-1,k}$).

In analysing these results, however, one must not lose sight of the fact that the two years during which the study was conducted were very different as to the amount of rainfall. While in 1992 the total precipitation during the vegetative period amounted to 488 mm, which is most favourable in terms of soybean development, 1992 was characterized by an extreme water deficit - only 212 mm (Figure 1).

Hence there were great differences in the relationships between the traits in question. In 1991, there were highly significant negative correlations between protein content and almost all of the traits under investigation, which is in agreement with the findings of various other authors (Johnson et al., 1955; Taira and Taira, 1971). However, Path-coefficient analysis showed that only oil content had a significant direct effect on grain protein content (Table 3).

Table 3 : Path coefficient analysis for protein content in 1991.

DIRECT EFFECTS	INDIRECT EFFECTS		
$p_{y1} = -0.073$	$r_{12}; p_{y2} = -0.490$*	$r_{13}; p_{y3} = -0.126$*	$r_{14}; p_{y4} = 0.017$
$p_{y2} = -0.695$*	$r_{21}; p_{y1} = -0.051$	$r_{23}; p_{y3} = -0.115$	$r_{24}; p_{y4} = -0.038$
$p_{y3} = -0.152$	$r_{31}; p_{y1} = -0.060$	$r_{32}; p_{y2} = -0.529$*	$r_{34}; p_{y4} = -0.021$
$p_{y4} = -0.123$	$r_{41}; p_{y1} = 0.010$	$r_{42}; p_{y2} = -0.213$*	$r_{43}; p_{y3} = -0.026$
$R^2_{y1234} = 0.832$			

Significantly at level 0.05 (*) and 0.01 (**)

The high negative correlation between yield and protein content was mostly a result of the the fact that yield is highly positively correlated with oil content - the direct effect of the trait on protein content was much smaller (-0.073).

There is a similar explanation for the high negative correlation between the duration of the vegetative period and protein content. The former trait is known to have a positive effect on yield (as confirmed in our study - 0.829), since it is during this period that the bulk of the vegetative mass, and, by virtue of this, of the yield itself, is formed (Patterson et al., 1977). Now, as both the duration of the vegetative period and yield are positively correlated with oil content, it is apparent that both of the traits must be negatively correlated with protein content.

The direct effect of the duration of the vegetative period on protein content was only -0.152. By contrast, the length of the reproductive period had no significant effect on protein content.

The following year, 1992, was most unfavorable for soybean due to both the total moisture deficit and the extremely unfavorable distribution of precipitation during the growing season.

The earlier-maturing genotypes suffered less from such conditions, producing yields that were somewhat higher than those of the late-maturing ones, in which all of the stages at which the formation of yield took place coincided with highly unfavorable weather conditions. Consequently, the relationships between the studied traits were different (Table 2).

Studies have shown that the largest portion of protein found in a soybean grain is synthesized in the period between 24 and 40 days after flowering as well as that protein synthesis peaks towards the end of this period (Rubel et al., 1972; Yazdi-Samadi et al., 1977).

In our study, those genotypes that flowered earlier also entered this stage earlier, thereby managing to partly avoid the negative effects of the drought that followed in July and August, a result of which was that they had somewhat higher yields and protein content than the late-maturing genotypes.

Both yield and protein content were negatively correlated with the duration of the vegetative period (-0.098 and -0.817, respectively), albeit there was also a negative correlation between these two traits themselves (-0.207).

Table 4 : Path coefficient analysis for protein content in 1992.

DIRECT EFFECTS	INDIRECT EFFECTS		
$p_{y1} = -0.195$	$r_{12}; p_{y2} = -0.208^{*}$	$r_{13}; p_{y3} = 0.073$	$r_{14}; p_{y4} = 0.123$
$p_{y2} = -0.268$	$r_{21}; p_{y1} = -0.151$	$r_{23}; p_{y3} = -0.148$	$r_{24}; p_{y4} = 0.127$
$p_{y3} = -0.743^{**}$	$r_{31}; p_{y1} = 0.019$	$r_{32}; p_{y2} = -0.054$	$r_{34}; p_{y4} = -0.039$
$p_{y4} = -0.216$	$r_{41}; p_{y1} = 0.111$	$r_{42}; p_{y2} = 0.158$	$r_{43}; p_{y3} = -0.134$
$R^2_{y1234} = 0.783$			

Significantly at level 0.05 (*) and 0.01 (**)

Path-coefficient analysis only confirmed the high significance of the negative effect of the duration of the vegetative period on grain protein content. Although negative, the direct effects of the other traits under investigation proved not to be significant (Table 4).

CONCLUSION

Based on our study, the following conclusions were made:

- All genotypes tested had a higher yield, a higher protein content, and a longer vegetative period in 1991 than in 1992.
- The highest yield was recorded in the line NS-L-2026 in 1991 (4.262 kg/ha) and the lowest in the line NS-L-1006 in 1992 (2.040 kg/ha), while the highest protein content was found in the line NS-L-82 in 1991 (43.9%) and the lowest in Šumadija in 1992 (36.4%).
- In 1991, a favorable year for soybean, protein content was highly correlated with yield, oil content, and the duration of the vegetative period.

Path-coefficient analysis demonstrated that only oil content had a significant direct effect on protein content in the grain.

- In 1992, which was an unfavorable year for soybean development, a high negative correlation was found between protein content and the duration of the vegetative period.

Path-coefficient analysis confirmed the high significance of the negative effect of this trait on grain protein content.

Our study has shown that, despite its high heritability, protein content primarily depends on environmental factors as well as that it is highly dependent on other traits, most notably oil content and the duration of the vegetative period.

REFERENCES

Brim C.A., 1973. Quantitative Genetics and Breeding. *In* Soybeans: Improvement, Production and Uses, Caldwell B.E., edit., Agronomy, 16, 155 - 186.

Brim C.A., Cockerham C.C., 1961. Inheritance of Quantitative Characters in Soybeans. *Crop Sci.*, 1, 187 - 190.

Burton J.W., 1987. Quantitative Genetics: Results Relevant to Soybean breeding. *Agron. J.*, 16, 211 - 247.

Burton J.W., Brim C.A., 1981. Recurrent Selection in Soybeans III. Selection for Increased Percent Oil in Seeds. *Crop Sci.*, 21, 31 - 34.

Erickson L.R., Voldeng H.D., Beversdorf W.D., 1981. Early Generation Selection for Protein in *Glycine max* x *Glycine soja* Crosses. *Can. J. Plant Sci.*, 61, 901 - 908.

Fehr W.R., Caviness C.E., 1977. Stages of Soybean Development. Special Report 80, Iowa State University, Ames, Iowa.

Hanson W.D., Probst A.H., Caldwell B.E., 1967. Evaluation of a Population of Soybean Genotypes with Implications for Improving Self - Pollinated Crops. *Crop Sci.*, 7, 99 - 103.

Hartwig E.E., Hinson K., 1972. Association Between Chemical Composition of Seed and Seed Yield of Soybeans. *Crop Sci.*, 12, 829 - 830.

Holbrook C.C., Burton J.W., Carter T.E., 1989. Evaluation of Recurrent Restricted Index Selection for Increasing Yield While Holding Seed Protein Constant in Soybean. *Crop Sci.*, 29, 324-329.

Hymowitz T., Collins F.I., Panczner J., Walker W.M., 1972. Relationship Between the Content of Oil, Protein, and Sugar in Soybean Seed. *Agron. J.*, 64, 613 - 616.

Ishige T., 1984. Biometrical Analysis and Estimation of the Number of Genes for Seed Protein Content of Soybean, *Glycine max* (L.) Merrill. *J. Agric. Res. Quart.*, 17, 230 - 235.

Johnson H.W., Robinson H.F., Comstock R.E., 1955. Genotypic and Phenotypic Correlations in Soybens and their Implications in Selection. *Agron. J.*, 47, 477 - 483.

Leffel R.C., 1988. High Protein Lines and Chemical Constituent Pricing in Soybean. *J. Prod. Agric.*, 2 (1), 111 - 115.

McKendry A.L., McVetty P.B.E., Voldeng H.D., 1985. Inheritance of Seed Protein and Seed Oil Content in Early Maturing Soybean. *Canadian Journal of Genetics and Cytology*, 27, 603 - 607.

Openshaw S.J., Hadley H.H., 1984. Selection indexes to Modify Protein Concentrations of Soybean Seeds. *Crop Sci.*, 24, 1 - 4.

Patterson D.T., Peet M.M., Bunce J.A., 1977. Effect of Photoperiod and Size of Flowering on Vegetative Growth and Seed Yield of Soybean. *Agron. J.*, 69, 631 - 635.

Rubel A., Rinne R.W., Canvin D.T., 1972. Protein, Oil, and Fatty Acid in Developing Soybean Seeds. *Crop Sci.*, 12, 739 - 741.

Sabbouh M.Y., Edwards L.H., 1985. Combining Ability for Seed Protein and Oil Content in Soybeans. *Soybean Genetics Newsletter*, 12, 99 - 100.

Sebern N.A., Lambert J.W., 1984. Effect of Stratification for Percent Protein in Two Soybean Populations. *Crop Sci.*, 24, 225 - 228.

Shannon J.G., Wilcox J.R., Probst A.H., 1972. Estimated Gains from Selection for Protein and Yield in the F_4 Generation of Six Soybean Populations. *Crop Sci.*, 12, 824 - 826.

Shorter R., Byth D.E., Mungomery V.E., 1976. Estimates of Selection Parameters Associated with

Protein and Oil Content of Soybean Seeds *Glycine max* (L.) Merr. *Aust. J. Agric. Res.*, 28, 211 - 222.

Simpson A.M., Jr., Wilcox, J.R., 1983. Genetic and Phenotypic Associations of Agronomic Characteristics in Four High Protein Soybean Populations. *Crop Sci.*, 23, 1077 - 1081.

Singh R.K., Chaudhary B.D., 1979. Biometrical Techniques in Genetics and Breeding. Kalyani Pub., Ludhiana, India, 39 - 80.

Taira Ha., Taira H., 1971. Influence of Location on the Chemical Composition of Soybean Seeds. *Nippon Shokuhin Kogyo Gakkaishi*, 1 (23), 6 - 12.

Taira Ha., Taira H., 1976. Effect of Size of Seed, Variety and Crop Year on the Chemical Composition of Soybean Seeds. *Nippon Shokuhin Kogyo Gakkaishi*, 40 (4), 181 - 196.

Thorne J.C., Fehr W.R., 1970. Incorporation of High Protein Exotic Germplasm into Soybean Populations by 2 - and 3 - way crosses. *Crop Sci.*, 10, 652 - 655.

Wehrmann V.K., Fehr W.R., Cianzio S.R., Çavins J.F., 1987. Transfer of High Seed Protein to High - Yielding Soybean Cultivars. *Crop Sci.*, 27, 927 - 931.

Yazdi - Samadi B., Rinne R.W., Seif R.D., 1977. Components of Developing Soybean Seeds, Oil, Protein, Sugars, Starch, Organic Acids, and Amino Acids. *Agron. J.*, 69, 481 - 486.