PLANT BREEDING AND SEED SCIENCE
Volume 45 Number 2 2001

Janusz Gołaszewski, Maria Idźkowska, Jadwiga Milewska, Irena Koczowska

Department of Plant Breeding and Seed Production, University of Warmia and Mazury in Olsztyn, Pl. Łódzki 3 10–724 Olsztyn

YIELD COMPONENT ANALYSIS WITH SYCA AND TDP IN FODDER PEA

ABSTRACT

The paper presents some theoretical assumptions of the SYCA (Sequential Yield Component Analysis) and the application of SYCA followed by TDP (Two–dimensional Partitioning) to analysis of the data from a plant breeding experiment with fodder pea. Partially balanced incomplete block design with 25 morphologically different breeding forms in 4 replications was applied. In both methods of data analysis plant height was the first trait in a sequence of independent traits, followed by different traits depending on the method.

The results of the analyses proved that in a morphologically highly differentiated population of fodder pea the contribution of plant height to the yield variability is reduced, with plant height to the first pod being one of the traits that have a significant effect on yield. According to the SYCA method, when the pea forms were divided into groups of plants similar in height, the effect of plant height as the first yield component was high and significant. Generally, the higher were the plants in the groups, the smaller was the share of the trait in the yield, although still relatively high and significant. For the purpose of yield component analysis in pea it is recommendable to divide the breeding material to groups of plants of a similar height.

As for the other yield components, the highest contribution into the final yield was attributed to the number of nodes with pods by plant height and seed weight by number of seeds calculated according to the SYCA and the number of nodes with pods calculated according to the TDP method, respectively.

The authors, who have used the two yield component analyses for several years, have gained enough experience to claim that the two methods can become effective statistical tools for the elaboration of yield components. Moreover, they can be useful not only in plant breeding studies but also in many other types of agricultural experimentation.

Key words: fodder pea, sequential yield component analysis, SYCA, TDP, two–dimensional partitioning, yield component analysis

INTRODUCTION

The main practical aim of a breeding programme is to obtain high yielding plant forms. At each stage of a breeding programme the breeder analyses the growth and development of plants and specifies significant traits, defined as yield components, which determine the yield volume. The relationship between yield and yield components can be interpreted

Communicated by Ludwik Spiss

using many statistical methods, of which the most common are correlation analysis, multiple regression and, recently, path analysis (e. g. Idźkowska *et al*. 1993a 1993b, Gołaszewski and Puzio–Idźkowska 1996, Boros *et al*. 1996). However, in all these methods the sequence of yield components, which is arranged in a mathematical model, seldom corresponds to the succession in which they differentiate at plant development stages. Obviously, it is not necessary to adhere to this succession in order to fulfil the formal conditions of analysis, but any diversion from the succession in which yield components are formed may lead to erroneous conclusions. This is usually due to the fact that, firstly, some variables introduced to the mathematical model are strongly correlated with each other or are derived from the recalculation of other variables present in the model, and secondly, each successive variable introduced to the model explains only part of the yield variability which has remained after the introduction of preceding cause variables. Consequently, if a cause variable strongly correlated to the cause variable previously incorporated in the model is introduced, its actual contribution to the formation of yield will be masked because the yield variability resulting from the mutual influence of the two traits has already been partially explained by the trait introduced earlier to the model.

The assumption of the sequential transfer of part of the total yield variability to successive traits differentiating during plant ontogenesis, which can be compared to the passing of a baton in a relay race, allows the researcher to capture an independent contribution of each trait present in the model to the final yield.

In this paper, Sequential Yield Component Analysis (SYCA) and Two–dimensional Partitioning (TDP) from analyses of variance for every yield component were applied to demonstrate a new opportunity of statistical evaluation of relationships between seed yield and its components in fodder pea. The assumptions on which the two methods are based and the relevant nomenclature were published by Eaton and Kyte (1978) and Eaton *et al*. (1986).

No examples for the application of SYCA can be found in Polish references. As regards TDP, the publications of Gołaszewski (1996) and Gołaszewski *et al*. (1996b, 1998) contain a complete computation algorithm of this method and its applications.

MATERIAL

The considerations are based on the results of measurements on fodder pea plants. The comparative experiments with pea (25 entries), using a partially balanced square lattice design with four replications, were conducted in 1998.

The pea plant material consisted of 15 original cultivars and 10 breeding strains of three–podded pea obtained by Puzio–Idźkowska (1997a, 1997b). Forty plants were sampled from each entry (10 plants per plot from each replication). The following measurements were taken: plant height, plant height to first pod, number of internodes to first pod, number of nodes with pods, number of pods per plant, number of seeds per plant, weight of 1000 seeds and seed weight per plant.

Statistical methods

Both methods, SYCA and TDP, include the orthogonalisation process which leads to the production of a set of non–correlated independent variables (Eaton *et al*. 1986, Gołaszewski 1996).

A starting point for the SYCA method is a multiplicative model, in which all the variables are assigned the same reference point (here: an individual plant) and each successive variable introduced to the model constitutes a proportional contribution attributed to the unit of the preceding variable.

Let us consider the algorithm of SYCA method applied to three independent variables X_1, X_2, X_3, which constitute a certain sequence, and a dependent variable Y.

Initially, the mathematical model is as follows:

$$X_1 \times \frac{X_2}{X_1} \times \frac{X_3}{X_2} \times \frac{Y}{X_3} = Y$$

When logarithms for both sides of the equation are found, the following additive model appears:

$$\log X1 + \log\left(X2 / X_1 \right) + \log\left(X_3 / X_2 \right) + \log\left(Y / X_3 \right) = \log Y$$

For the purpose of simplification, let us assume that:

$$X_1^{'} = \log(X_1); \quad X_2^{'} = \log(X_2 / X_1); X_3^{'} = \log(X_3 / X_2);$$
$$X_4^{'} = \log(Y / X_3); \quad Y^{'} = \log(Y)$$

Orthogonalisation. A new set of non–correlated independent variables is created corresponding to the input variables:

$$X_1^{''} = X_1^{'}$$
$$X_2^{''} = X_2^{'} - E\left(X_2^{'}\right),\ and\ E\left(X_2^{'}\right) = b_0 + b_1 \times X_1^{'}$$
$$X_3^{''} = E\left(X_3^{'}\right),\ and\ E\left(X_3^{'}\right) = b_0 + b_1 \times X_1^{'} + b_2 \times X_2^{'}$$
$$X_4^{''} = E\left(X_4^{'}\right),\ and\ E\left(X_4^{'}\right) = b_0 + b_1 \times X_1^{'} + b_2 \times X_2^{'} + b_3 \times X_3^{'}$$

Determination of the per cent contribution of each successive independent variable into the resulting trait.

a) when the variables are standardised

$$E(Y') = b_0 + b_1 \times X_1'' + b_2 \times X_2'' + b_3 \times X_3'' + b_4 \times X_4''$$

where

$$r_i^2 = b_i^2$$

and

$$r_1^2 + r_2^2 + r_3^2 + r_4^2 = 1$$

b) when the variables are not standardised, contribution of each cause trait into the resultant one may be determined as the square of simple correlation coefficients between the variables obtained in the orthogonalisation process and the resultant variable.

TDP may be viewed as an extension of SYCA if analyses of variance are carried out for the variables transformed as above and the results are incorporated in a two–dimensional table (Eaton *et al*. 1986). However, the fact that the results from the SYCA on which the conclusions are formed are indices rather than actual traits could be a problem here; besides, they are presented in a logarithmic scale. On the other hand, the indices have their practical meaning, for example the relation: number of seeds divided by number of pods describes the average number of seeds per pod.

The procedure described above for SYCA method can also be applied to the original data X_1, X_2, X_3 and Y, starting the analysis by first creating the additive model and then applying the orthogonalisation process. The determination coefficients r^2 set according to this method will correspond to the successive participation of non–correlated independent variables and, in comparison with the TDP, they will correspond to the total sums of squares from analyses of variance.

In this paper analyses of data for fodder pea were conducted according to SYCA with four independent variables: plant height, number of nodes with pods, number of pods, number of seeds and the dependent variable seed weight. Besides, the TDP method accounted for all the traits of pea described was applied.

RESULTS

Table 1 contains the data on the contribution of pea yield components in per cent of yield variation determined according to SYCA. Groups of plants of different height, both original cultivars and strains, were included.

The results of the SYCA analysis for groups of plants different in height show that all the yield components taken into consideration had a significant influence on the seed yield. The dominant effect was produced by the plant height (PH) and the proportion which determines the number of podded nodes per unit of plant height (PN/PH), that is the first two traits introduced to the model. The average contribution of these two components to the yield was about 60%. It is interesting to no–

tice that although the total contribution of the two components was similar, the proportion between them changed at different lengths of plants. The contribution of the former variable was declining while the plant height was increasing; at the same time the proportional contribution of number of nodes with pods per plant height was on the increase. This tendency was observed for both cultivars and strains. As regards the other variables, i. e. the number of pods per node (NP/PN), number of seeds per pod (NS/NP) and the weight of one seed (SW/NS), it is difficult to determine unambiguously their relation to the grouping of plants. It may be assumed that each of these variables had a significant effect on the yield of seeds per plant and contributed similar information.

An analogous analysis on mean values of traits from all the experimental plots (100) confirmed a high contribution of the PN/PH component. The effect of the plant height turned out to be rather small, while the other traits had a much higher contribution to the yield than the contribution analysed in groups according to plant height; the later the trait developed in the plant ontogenesis, the higher its contribution to the yield.

Total sums of squares from analyses of variance and percentage contribution of components to pea seed yield according to TDP are presented in Table 2. Two sources of variability were distinguished – entries and environment related. The term „environmental variability" includes the variability between replications, between blocks and respective genetic x environmental interaction which compose the experimental error of the incomplete blocks design. The pea experiment was established on a highly variable soil, and the weather in 1998 was generally favourable to the vegetation of plants (Szczepańska 1999).

The two traits, plant height and number of internodes to the first pod, had a non–significant effect on the total variability of pea seed yield. The contribution of the plant height, 1%, was the same as determined with SYCA on the mean values for plots (Table.1). This is connected to the fact that the plant height was the first trait taken into consideration and constituted a reference point for the remaining traits (independent ones and the dependent one). The studies conducted earlier by Gołaszewski *et al.* (1998) with the TDP method applied to the analysis of 15 pea entries, including 12 medium–height (about 80 cm) breeding strains of seed type fodder pea showed that in a population of pea uniform in terms of height the contribution of this trait to the yield was approximately 30%. These observations are in agreement with the results produced according to SYCA, obtained on pea strains of a similar plant height (Table1).

Plant height to the first pod, commonly considered as indicative of plant earliness, was the only morphological trait which significantly determined the yielding of pea plants. It is also interesting to notice that the contribution of the entry variability for plant height and plant height to the first pod to the total variability was higher than that of the environmental variability. For the other traits the relation between the two terms of variability was opposite.

Table 1

Contribution of pea yield component [%] according to SYCA.

Mean height [cm]	No of plants	Plant height (PH)	No of nodes with pods/ Plant height (PN/PH)	No of pods/ No of nodes with pods (NP/PN)	No of seeds/ No of pods (NS/NP)	Seed weight/ No of seeds (SW/NS)	Seed weight (SW)
				Original cultivars[1]			
69	200	41	19	12	14	14	100
103	120	38	29	18	6	9	100
167	280	27	38	8	13	14	100
				Strains[1]			
67	200	39	18	12	20	11	100
157	200	20	36	14	14	16	100
Average for pea[2]		1	30	17	22	30	100

[1]analyses for traits recorded from single plants (n – number of plants)
[2]analyses for averages per plot

Table 2.

Two–dimensional partitioning of pea yield variation.

Parameters	Sums of squares from analyses of variance			Percentage of the total yield variation		
	Variability			Variability		
Variable (trait)	Entries	Environ–mental	Total[a]	Entries	Environ–mental[b]	Total
Plant height	6.3**	0.4	6.7	1	0	1
Plant height to 1st pod	37.2**	14.5	51.7*	3	1	4
No of internodes to 1st pod	6.8*	12.0	18.8	1	1	2
No of nodes with pods	124.1	205.4	329.5**	11	18	29
No of pods	74.0**	103.2	177.3**	6	10	16
No of seeds	124.6**	148.5	273.0**	11	13	24
1000 seed weight	105.4**	118.7	224.1**	9	10	19
Residual variability	14.1	46.3	60.3*	1	4	5
XX[c]	–144.5	144.5		–13	13	
Seed weight	347.9**	793.5	1141.4	30	70	100

[a] – the totals were tested by regression of noncorrelated independent variables and yield
[b] – zero results from rounding
[c] – cross–products – possible interaction between components calculated for each source of varia–tion in analysis of variance; the sum of cross–products equals zero
*, ** – significant at P=0.05 and P=0.01, respectively

The total variability of the other pea yield components: number of nodes with pods, number of pods, number of seeds and weight of one thousand seeds, was significant, with its contribution to the final yield at 88%. These traits were strongly modified by the environment, which is confirmed by a higher environmental effect compared to the treatment effect. A considerably high contribution to the total environmental variability was ascribed to the number of nodes with pods. No significant differences between the entries were found out for this trait, although it was a dominant yield component and the experimental material was highly variable.

DISCUSSION

When all the pea forms were included into a single analysis, the results from the two methods (SYCA and TDP) lead to the same conclusion. Thus, it can be assumed that in screening experiments with pea forms highly differentiated in their height (as in our experiment) the effect of plant height as an initial trait in yield component analysis was negligible and the importance of the other yield components was significant. However, dividing such a population into subpopulations of pea plants similar in height revealed a high contribution of the plant height to the final yield (SYCA). Another interesting observation from SYCA analysis is that in groups of plants differentiated in height, the share of height in the yield was decreasing for higher plants. The results from the SYCA agree with the results obtained in the authors' earlier studies, where pea forms similar in height were analysed in a field experiment (Gołaszewski *et al*. 1998).

The analysis of the pea yield components carried out according to TDP showed a very small effect of the height of plants on yield, but suggested that height of plants to the first pod was one of the morphological traits which were significant yield components in our studies. Strong, significant entry effect of the plant height to the first pod together with a small environmental effect may suggest that the trait should be suitable not only for the determination of the earliness of plants, but also for selection of high yielding plants in breeding tests on a wide spectrum of pea plant morphotypes.

The other significant yield components: number of nodes with pods, number of pods, number of seeds and 1000 seed weight, even so their shares in the yield are very high, depend to the great extent on environmental conditions. Thus, the appearance of the traits and eventually the yield is strongly modified by the environment. For these traits, the TDP method showed higher environmental variability than entry variability.

Concluding, it seems advisable to divide pea forms into groups of plants approximately similar in height when evaluating pea yield components and yield through field experiment. What follows is that entries for tests must be appropriately located on the field in the sense of their spatial distribution. It bears a new experimental situation when the experiment can consist incomplete blocks of entries similar in height with different capacity of the blocks. It also suggests some modification of the analysis of variance by including contrast in the entry variation or an additional source of variation, i. e. morphologically different groups of pea forms.

Having applied the two yield component analyses for several years, the authors feel confident enough to claim that both of them can become effective statistical tools for the elaboration of yield components. Moreover, the methods can be useful not only in plant breeding studies but also in many types of agricultural experimentation. The plant breeder will find the results obtained with SYCA and TDP of scientific value because:

1. Increments in r^2 measure the contribution of every new independent variable when it is added in succession to the yield equation, in the case of SYCA – to sum up to 100% of yield variation. Variables in the SYCA method are indices, although they have practical value, e. g. "weight of seeds/number of seeds" is practically the weight of one seed, that is a trait similar to 1000 seed weight but recalculated, not obtained empirically. The SYCA and TDP methods converge when the calculation is made on original data for traits, but here to sum up to 100% an extra trait, residual variability, is created.
2. Thanks to two–dimensional partitioning (TDP) of the total yield variation according to the successive contribution of independent variables – one direction, and the sources of variation from the proper analyses of variance completed for all the traits studied – the other direction, concise presentation and easy interpretation of the results are attainable. In this type of presentation total yield variation can be analysed not only as the effect of each independent trait studied but also as a result of different sources of variation calculated for each trait and derived from the analysis of variance applied.
3. The TDP method can be used in synthesis of the data from many locations (years), in which case the method enables the breeder to assess part of the total yield variability of each independent trait attributed to environmental factors, experimental factors (i. e. varieties, breeding forms, etc.) and to the environment × treatment interaction.

REFERENCES

Boros L., Sawicki J., Gołaszewski J. 1996. Path–coefficient analysis of seed yield components in soybean (*Glycine max* L. Merr.) and dry bean (*Phaseolus vulgaris* L.). XXVI International Biometrical Colloquium, Abstracts, 10–11.

Eaton G. W., Bowen P. A., Jolliffe P. A. (1986). Two–dimensional partitioning of yield variation. Hortscience. 21, 1052–1053.

Eaton G. W., Kyte T. R. 1978. Yield component analysis in the cranberry. J. Am. Soc. Hort. Sc., 103: 578–583.

Gołaszewski J. 1996. A method of yield component analysis. Biometrical Letters. 33: 79–88.

Gołaszewski J., Idźkowska M., Milewska J. 1998. The TDP method of yield component analysis in grain legume breeding. Journal of Applied Genetics. 39 (4): 299–308.

Gołaszewski J., Milewska J., Koczowska I. 1996. Dwukierunkowy rozkład sum kwadratów w analizie zmienności plonowania rodów bobiku. Acta Acad. Agricult. Tech. Olst., 64: 147–153.

Gołaszewski J., Puzio–Idźkowska M. 1996. Analiza ścieżkowa w ocenie współzależności niektórych komponentów plonu nasion grochopeluszki. Biul. IHAR, 200: 317–322.

Idźkowska M., Koczowska I., Gołaszewski J., Grabowski St. 1993a. Analiza współczynników ścieżek u żyta. Cz. I. Analiza współczynników ścieżek w ocenie współzależności cech determinujących masę ziarna z kłosa". Acta Acad. Agricult. Tech. Olst., 56: 17–24

Idźkowska M., Gołaszewski J., Koczowska I., Grabowski St. 1993b. Analiza współczynników ścieżek u żyta. Cz. II. Analiza współczynników ścieżek w ocenie współzależności cech determinujących masę ziarna z rośliny. Acta Acad. Agricult. Tech. Olst., 56: 25–30.

Jolliffe P. A., Eaton G. W., Lovett Doust J. 1982. Sequential analysis of plant growth. New Phytologist, 92: 287–296.

Puzio–Idźkowska M. 1997a. Grochopeluszka trójstrąkowa – nowa forma rośliny strączkowej typu nasiennego. Zesz. Prob. Post. Nauk Rol., 439:49–53.

Puzio–Idźkowska M. 1997b. Nowy genotyp grochu pastewnego. Zesz. Prob. Post. Nauk Rol., 446: 113–118.

Szczepańska A. 1999. Wpływ zmienności glebowej i oddziaływań sąsiedzkich na efektywność porównań międzyobiektowych w doświadczeniu hodowlanym z grochem siewnym. Praca magisterska, Biblioteka Uniwersytetu Warmińsko–Mazurskiego w Olsztynie