

# The TDP method of seed yield component analysis in grain legume breeding

Janusz GOŁASZEWSKI, Maria IDŹKOWSKA, Jadwiga MILEWSKA

Department of Plant Breeding and Seed Production, Olsztyn University of Agriculture  
and Technology, Poland

**Abstract.** The results of plant breeding trials with populations of fodder pea strains and broad bean hybrids were the basis of consideration on the interrelationship between some traits – the yield structure elements. Developed by Eaton, a relatively new method of yield component analysis called the two-dimensional partitioning method (TDP) was applied to analyse the data. The method, which combines multiple regression and ANOVA, allows for concise tabular presentation and simple interpretation of the distribution of traits in one direction and the sources of variance according to ANOVA model in the other direction. Additionally, the interpretation of the results was supported by such standard statistical techniques as ANOVA, simple and multiple regression and path analysis. The main components of pea yielding were plant height and the number of pods per plant. Among the analysed characters of broad bean the number of nodes with pods on the main stem, which turned out to be the determinant of broad bean yielding, might be strongly affected by environmental conditions. The number of nodes with pods might be considered a selecting character of high potential yielding of broad bean genotypes.

**Key words:** ANOVA, broad bean, fodder pea, multiple regression, path analysis, *Pisum sativum*, TDP method, *Vicia faba*, yield components.

## Introduction

The main aim of plant breeding field trials with new cultivars is to evaluate the variability of several characters, especially the so-called yield components that determine the productivity of a crop. It is common analytical practice in the studies to apply the analysis of variance (ANOVA) in order to compare

Received: May 1998.

Correspondence: J. GOŁASZEWSKI, Department of Plant Breeding and Seed Production, Olsztyn University of Agriculture and Technology, Plac Łódzki 3, 10-724 Olsztyn, Poland.



C-21103

treatment differences and the multivariate method (i.e. multiple correlation and regression) to assess the interrelationships between traits. Since these methods are used separately, conclusions tend to be limited to a particular trait. In all the methods mentioned above, it is assumed that an analysed trait acted alone and is excluded from the set of other traits. Although multiple regression embraces some variables, each trait included into the additive model is treated separately in the sense that the relationship between each independent trait and the dependent one (regression coefficient) is assessed in the absence of the other independent variables; the total measurement gives information about the relation between the set of independent traits and the dependent trait. An extended interpretation of the multiple regression results leads to path analysis, where total variation of relation between an independent trait and the dependent trait is divided into a direct effect describing a pure relationship between the two traits and indirect effects by the other independent traits.

In reality, however, characters develop at successive stages of ontogenesis and thus they should be treated as a sequence of dependencies, where each character contributes to the one(s) to follow. By reversing this idea, it may be said that each character depends on a sequence of traits which appeared earlier in plant growth.

The statistical method presented here (i.e., two-dimensional partitioning – TDP), developed by EATON (1986), is a compilation of ANOVA and multiple regression analysis. It is based on a transformation of original data, which assumes that the variation of preceding characters is excluded from the total variation of each successive character and the contribution of a character into yield carries pure information of itself only. Eventually, the method enables to present the results in a concise, two-dimensional table, as the incremental contribution of yield components into the final character – yield (one direction) and the division of that contribution into parts related to the sources of variation of the appropriate ANOVA (second direction). Statistical assumptions of the transformation and the calculation procedure can be found in the works of WINER (1971) and GOŁASZEWSKI (1996).

The purpose of the study was (i) to evaluate some seed yield components, which determine the yielding of fodder pea (*Pisum sativum* L.) and broad bean (*Vicia faba* L.) breeding materials in north-eastern Poland and (ii) to evaluate a relatively new method of yield component analysis (Eatons TDP) versus the statistical techniques traditionally applied for analysis of plant breeding materials.

## Material and methods

The study was based on the results of two field experiments carried out in 1996: one with fodder pea, the other with broad bean. Both used a randomized block design (RBD) with 3 replications for fodder pea and 4 for broad bean.

The studied materials were plants of medium height (about 80 cm): 12 strains and 3 cultivars (Pegro, Grapis, Pelikan) of fodder pea; and 14 hybrids and 2 Polish cultivars (Nadwiślański, Dino) of broad bean. The breeding materials were obtained at the University of Agriculture and Technology in Olsztyn, Poland (PUZIO-IDŹKOWSKA 1995, MILEWSKA et al. 1996).

After harvest, the following traits were measured: plant height, total number of nodes, number of branches, number of pods, number of seeds, 1000-seed weight and seed yield in the case of fodder pea; and plant height, number of fruiting nodes on the main stem, number of pods, number of seeds, 1000-seed weight and seed yield in the case of broad bean. Except the number of fruiting nodes on the main stem recorded for broad bean, all the other measurements were taken on a per plant basis. It can be assumed that the sequence of the characters reflects their differentiation at the consecutive stages of plant development.

The data were analysed using ANOVA, simple and multiple regression extended to path coefficient analysis and Eaton's two-dimensional partitioning (TDP) procedure. The main steps in the computational procedure of the TDP method, so far used only occasionally, are as follows:

1. Set the original characters in a sequential order of their appearance at the successive stages of their differentiation  $V_1, V_2, \dots, V_n$ , where  $n$  is the number of characters, and now  $V_n = Y$  is the yield.

2. Transform the characters to a new set of uncorrelated characters  $X_1, X_2, \dots, X_n$ ; the first character ( $X_1$ ) and yield ( $Y$ ) after transformation will be the same as before, and a new extra variable ( $X_n$ ) will embrace residual variability:

$$\begin{aligned}
 \hat{V}_2 &= b_{2,0} + b_{2,1}V_1 \\
 \hat{V}_3 &= b_{3,0} + b_{3,1}V_1 + b_{3,2}V_2 \\
 &\vdots \\
 &\vdots \\
 \hat{V}_n &= b_{n,0} + b_{n,1}V_1 + \dots + b_{n,n-1}V_{n-1}
 \end{aligned}
 \tag{1}$$

and a new set:

$$\begin{aligned} X_1 &= V_1 \\ X_2 &= V_2 - \hat{V}_2 \\ &\vdots \\ &\vdots \\ X_n &= V_n - \hat{V}_n \text{ (an extra variable)} \end{aligned} \quad (2)$$

3. For a new set of data calculate multiple regression to obtain coefficients of regression for further scaling

$$\hat{Y} = b_0 + b_1 X_1 + \dots + b_n X_n \quad (3)$$

and simple regression between new variables and yield to measure the incremental contribution of successive characters to yield (coefficients of determination).

In the case of standardization of the new set of variables

$$(X' = \frac{X - \bar{X}}{\text{standard deviation}})$$

$$\hat{Y} = b_0 + b_1 X'_1 + \dots + b_n X'_n \quad (4)$$

and coefficients of determination of the character  $i$  are calculated directly from (4) as  $r_i^2 = b_i^2$ .

4. Scale variables to units of  $Y$  by multiplying each variable by an appropriate coefficient from [3] and for the variables prepared in this way calculate ANOVA. For the following sources of variation specified in ANOVA the sum of sums of squares of yield components is equal to the proper sum of squares for yield.

5. Create a two-dimensional table with the results expressed as sums of squares or percentages.

## Results

### Fodder pea

Simple correlations and path coefficients for fodder pea are presented in Table 1. In the set of independent variables all possible correlations were significant, except the relationship between the number of pods per plant and 1000-seed weight. Besides, the correlations between 1000-seed weight and the other yield components were negative. The strongest phenotypic correlation with yield was demonstrated by the number of seeds per plant ( $r = 0.836$ ) and

**Table 1.** Simple correlations and path coefficients for pea yield components

Variables – Traits (original data)	Plant height (cm) V <sub>1</sub>	Nodes/plant V <sub>2</sub>	Branches/ plant V <sub>3</sub>	Pods/plant V <sub>4</sub>	Seeds/plant V <sub>5</sub>	1000-seed weight (g) V <sub>6</sub>
Correlation coefficients ( $r_{tab}(P=0.05)=-0.290$ , $r_{tab}(P=0.01)=0.375$ )						
V <sub>1</sub>	1					
V <sub>2</sub>	0.857	1				
V <sub>3</sub>	0.358	0.392	1			
V <sub>4</sub>	0.492	0.511	0.680	1		
V <sub>5</sub>	0.685	0.684	0.672	0.859	1	
V <sub>6</sub>	-0.405	-0.382	-0.332	-0.246	-0.483	1
Path coefficients (random factor $P_e=0.351$ , underlined direct effects)						
V <sub>1</sub>	<u>-0.076</u>	-0.065	-0.027	-0.037	-0.052	0.031
V <sub>2</sub>	0.080	<u>0.093</u>	0.036	0.048	0.064	-0.036
V <sub>3</sub>	0.002	0.002	<u>0.005</u>	0.003	0.003	-0.002
V <sub>4</sub>	-0.044	-0.046	-0.061	<u>-0.090</u>	-0.077	0.022
V <sub>5</sub>	0.779	0.778	0.764	0.977	<u>1.137**</u>	-0.549
V <sub>6</sub>	-2.201	-0.189	-0.164	-0.122	-0.239	<u>0.496**</u>
Phenotypic correlation with yield (Y)	0.540	0.573	0.553	0.779	0.826	-0.038
R <sup>2</sup> (%)	29.2	32.8	30.6	60.7	69.9	0.1

\*\* significant at P = 0.01

the number of pods per plant ( $r = 0.779$ ). Only the simple correlation of yield with 1000-seed weight was non-significant. The analysis of direct and indirect path coefficients permits to examine more closely the interrelationships in the set of studied characters. It was found that only the number of pods per plant and 1000-seed weight had a significant direct effect on yield. Furthermore, the two characters masked effects of plant height and the other yield components. In spite of the significant correlations with yield, their direct effects were hardly present. In such cases the magnitude of correlation coefficients was caused by a strong positive indirect effect of the number of seeds per plant and a relatively high negative indirect effect of 1000-seed weight. Exactly the opposite effect was displayed by 1000-seed weight. Lack of correlation with yield was accompanied by a high positive direct effect, which was neutralized by a nearly equal negative indirect effect of the number of seeds per plant. Similar conclusions were drawn from the earlier study on edible field pea by GOŁASZEWSKI and PUZIO-IDŹKOWSKA (1996). The number of pods per plant, the average number of seeds per plant and 1000-seed weight

**Table 2.** Two-dimensional partitioning of pea yield variation (%)

Sources of variation	Transformed variables								
	plant height (cm)	nodes/plant	branches/plant	Pods/plant	seeds/plant	1000-seed weight (g)	residual variability	cross-products <sup>a</sup>	seed yield (g)
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	XX	Y
Blocks	2.0**	0.1	1.3	1.1	0.9*	1.7*	0.3	2.0	9.5*
Treatments	25.9**	3.5*	4.8	9.4*	3.5*	9.0*	4.1	2.2	62.4*
Error	1.3	1.0	6.3	8.4	2.4	5.2	7.8	-4.2	28.2
Total = R <sup>2</sup>	29.2**	4.6	12.4*	18.9*	6.8	15.9*	12.2*		100.0

\*, \*\* significant at  $P = 0.05$  and  $P = 0.01$ , respectively

<sup>a</sup> – embrace all possible interactions between treatments and component pairs

had predominantly positive direct effects. The latter acted antagonistically in relation to the other analysed yield components.

Two-dimensional partitioning of sums of squares (in %) is showed in Table 2. The results reveal another structure in the set of characters – yield components and yield. Among the characters differentiated at the successive stages of plant development plant height had the greatest impact on yield. Besides, a significant independent effect on yield was produced by the number of pods per plant (18.8%), 1000-seed weight (15.9%) and the number of branches (12.4%). A relatively high random factor ( $P_e = 0.351$ ) and its significant effect (12.3%) indicate the possibility of including more character(s) in the model.

The main constituent of the total sum of squares for yield was treatment variation. One third of the total yield variation may be attributed to environmental variability (blocks + error). Except the number of branches per plant, the treatment variation for the other characters was significant. It is worth noticing that treatment variation was significant for all the original variables.

### Broad bean

The interrelationships among broad bean yield components were quite different than in fodder pea (Table 3). The number of fruiting nodes on the main stem correlated significantly with the number of pods per plant and the number of seeds per plant. A strong correlation was observed for the number of pods and seeds per plant. These characters were also correlated strongly with yield. The correlations of yield with plant height and 1000-seed weight were non-significant. Conclusions similar to those suggested for fodder pea can be drawn

**Table 3.** Simple correlations and path coefficients for faba bean yield components

Variables – Traits (original data)	Plant height (cm)	Fruiting nodes/main stem	Pods/plant	Seeds/plant	1000-seed weight (g)
	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>	V <sub>4</sub>	V <sub>5</sub>
Correlation coefficients ( $r_{tab(P=0.05)}=0.248$ , $r_{tab(P=0.01)}=0.323$ )					
V <sub>1</sub>	1				
V <sub>2</sub>	0.201	1			
V <sub>3</sub>	0.007	0.853	1		
V <sub>4</sub>	0.182	0.844	0.956	1	
V <sub>5</sub>	-0.461	-0.056	-0.129	-0.234	1
Path coefficients (random factor $P_e=0.036$ , underlined direct effects)					
V <sub>1</sub>	<u>0.005</u>	0.001	0.000	0.001	-0.002
V <sub>2</sub>	0.005	<u>0.023</u>	0.020	0.019	-0.001
V <sub>3</sub>	0.000	-0.008	<u>-0.010</u>	-0.009	0.001
V <sub>4</sub>	0.184	0.857	0.971	<u>1.016</u> **	-0.238
V <sub>5</sub>	-0.126	-0.015	-0.035	-0.064	<u>0.274</u> **
Phenotypic correlation with yield (Y)	0.068	0.858	0.946	0.963	0.034
R <sup>2</sup> (%)	0.5	73.6	89.5	92.7	0.1

\*\*significant at P = 0.01

from the analysis of the path coefficients. Highly significant direct effects on yield are produced by the number of seeds per plant and 1000-seed weight. The high correlation values of these characters with yield resulted mainly from strong indirect effects of the number of seeds per plant. The indirect effects for 1000-seed weight, although also negative, were not so distinctively ex-

**Table 4.** Two-dimensional partitioning of faba bean yield variation (%)

Sources of variation	Transformed variables							
	plant height (cm)	fruiting nodes/main stem	Pods/plant	seeds/plant	1000-seed weight (g)	residual variability	cross-products <sup>a</sup>	seed yield/plant (g)
	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	XX	Y
Blocks	0.1**	22.2**	0.1	0.2	0.2	0.0	-2.6	20.2**
Treatments	0.2*	14.7	6.3*	1.8*	2.1*	0.0	-8.4	16.7
Error	0.2	37.3	9.4	2.3	2.8	0.1	11.0	63.1
Total = R <sup>2</sup>	0.5	74.2**	15.8**	4.3	5.1	0.1		100.0

\*, \*\* significant at P = 0.05 and P = 0.01, respectively

<sup>a</sup> – embrace all possible interactions between treatments and component pairs

pressed as in fodder pea. The random factor was minuscule ( $P_e = 0.036$ ), which can indicate an optimal choice of the characters.

The analysis of the results with TDP (Table 4) shows clearly that the main yield component was the number of fruiting nodes on the main stem. This character, together with the number of pods per plant, accounted for 90% of total yield variability. Moreover, it is worth noticing that 80% of its variation could be attributed to environmental factors.

Unstable yielding and high susceptibility to environmental conditions of this species was confirmed by high values for blocks and error variation (about 85%). Before the data was transformed, a significant treatment variation in plant height and the number of seeds per plant was observed. After the transformation the treatment variation in the number of pods and seeds per plant and 1000-seed weight appeared also significant.

## Discussion

Fodder pea and broad bean play a major role in fodder production in north-eastern Poland; broad bean is commonly called there the Soya Bean of the North. Therefore, one of the main tasks of the breeder is to increase the seed yield and to improve its stability.

Plant height and yield components of the two species analysed in the paper showed their distinctly different effects on yielding. The main character that determined final yield of fodder pea was plant height. Indirectly, this character also determined the formation of the other yield components. Among the yield components of fodder pea, the most affected by environmental factors was the number of branches per plant (61%) while the least affected was plant height (11%) and the total number of nodes (24%). The variability of the number of pods and seeds was almost equally divided between environmental and treatment (genetic) factors.

In broad bean the main contributor to the final yield was the number of fruiting nodes on the main stem, but also this character was the most remarkably affected by environmental factors – about 30% of blocks and about 50% of error variation. In comparison with fodder pea, the treatment variation of all the broad bean yield components was relatively small (20-40%). This proves that the environment has a significant impact on the formation of successive characters – broad bean yield components.

Application of the two-dimensional partitioning method in the experiments with fodder pea and broad bean enabled the researcher to identify the characters which had the most significant effect on yield. These were: plant height for



fodder pea strains and the number of fruiting nodes on the main stem for broad bean hybrids. Both characters are easy to measure at early stages of plant development and could be considered as selection characters in breeding programmes with those materials. The choice of the number of fruiting nodes in broad bean is in accordance with the breeders' expectations and it is also suggested in literature (e.g., POULSEN 1974, GOŁASZEWSKI et al. 1996, MIŁEWSKA et al. 1996), but in fodder pea the choice of plant height as a selection character seems to be hazardous. Pea plants with long stems bear seed later than those with shorter stems, and their harvest is usually much more difficult. Thus, plant height should not be a top selection character in the breeding of high yielding fodder pea. However, the results of our study may be a premise for breeders that high yielding fodder pea plants should not be too short. Another important character in the sequence of yield components was the number of pods per plant, which is definitely a very good index of high pea yielding.

In the light of the TDP method, strong determination of fodder pea yield by plant height may be interrelated with the fact that the character is the first one in the model. It remains unchanged during the transformation process and, being the first character in the model, it is not dependent on the others. In reality, plant height in pea is also conditioned by many genetic, physiological and environmental factors.

In comparison with the standard methods of analysis of plant breeding data such as ANOVA, simple and multiple regression and path coefficients, the TDP method contributes additional valuable information, which gives a possibility of broader interpretation of the results of breeding trials.

## Conclusions

1. Plant height and the number of pods per plant in fodder pea strains and the total number of fruiting nodes on the main stem in broad bean hybrids were the main components of high yielding of these populations in north-eastern Poland in 1996.

2. The two-dimensional partitioning method reveals a new structure of interrelationships in the set of yield components and can be used to detect the essential characters determining the yielding of a breeding population. A simple computational procedure and comprehensible results of the analysis make a good premise for further applications of the method. A collection of many examples of analysis of plant breeding data with TDP will probably allow to assess fully the usefulness of the method.

**Acknowledgements.** This study was supported by The State Committee for Scientific Research Project No. 5 PO6A 00711.

#### REFERENCES

- EATON G.W. (1986). Two-dimensional partitioning of yield variation. Hort. Sci. 21(4): 1052-1053.
- GOŁASZEWSKI J. (1996). A method of yield component analysis. Listy Biometryczne – Biometrical Letters 33(2): 79-88.
- GOŁASZEWSKI J., PUZIO-IDŹKOWSKA M. (1996). Path analysis in estimation of interdependence of some seed yield components of edible field pea. (in Polish). Bulletin of Plant Breeding and Acclimatization Institute 200: 317-322.
- GOŁASZEWSKI J., MILEWSKA J., KOCZOWSKA I. (1996). Two-dimensional partitioning in yield variation of field bean (*Vicia faba* L.) strains. (in Polish). Acta Acad. Agricult. Tech. Olst., Agricultura 63: 147-153.
- MILEWSKA J., GOŁASZEWSKI J., KOCZOWSKA I. (1996). The variability of some morphological traits of broad bean strains. (in Polish). Bulletin of Plant Breeding and Acclimatization Institute 200: 339-348.
- POULSEN M.H. (1974). Effect of mode seed setting on yield and earliness in *Vicia faba*. Z. Pflanzenzüchtung 72(2): 120-131.
- PUZIO-IDŹKOWSKA M. (1995). An attempt of multipodded field pea (*Pisum sativum* L.) breeding. Proc. 2nd Eur. Conf. Grain Legumes, Copenhagen – Denmark: 217.
- WINER B.J. (1971). Statistical principles in experimental design. (2nd edn.). McGraw-Hill Book Company, New York.