T1988PRO86

THE EXAMINATION OF VARIABLES INFLUENCING THE YIELD OF SUNFLOWER WITH DISCRIMINANCE ANALYSIS

E.Baráth, K.Simits, M.M.Omed
University of Agricultural Sciences H-2103 Gödöllő Páter K.u.1.
Hungary

On the basis of data coming from agricultural fields the importance of a given factor on the yield of sunflower was studied taking into cosideration the complex effect of several factors. A multivariate method - the discriminance analysis /DA/-was used.

In the example data of 170 fields - 12.000 hectars - of NSH 26 hybrid /previous crop was winter wheat/ were analysed. The cropping system and its conditions /soil, precipitation, etc/ were discribed with 18 variables.

Two groups were formed: fields with low /under the mean/ and fields with high /above the mean/ yield. The division of the multivariate distance of DA - the D^2 value - shows the effect of a given variable in % according to the classification into yield groups.

Concerning to the results of the discriminance analysis the most important variables were the "realised yield %" /harvested/planned yeald t/ha/ and the "yield of the previous crop". The total rainfall, the texture of the soil and its high K_2O ppm contents were also important.

The negative effect of late sowing can be proved as an important indirect effect on the yield. The significant importance of the required germination is proved by the effect of the "realised plant number %" /number of harvested/planned plants/, as well. In this analysis the phosphate contents of the soil hadn't any effect on the yield because all the fields were well supplied with P.

Finally, it can be seen that the discriminance analysis is an adequate method to analyse the importance of a given factor in multivariate system in agriculture.

INTRODUCTION

The actual yield of an agricultural field is influenced by several factors. In the same year the yields reached in different fields may show even a deviation of 2,0 t/ha.

The question is raised how important is the effect of a given factor of the cropping system on the yield of sunflower taking into consideration the complex effect of several factors.

The complex effects can't be examined in experiments because of the big number of variables. That's why we deal with scientific evaluation of sunflower field data coming from the production.

In the paper the importance of 18 variables is analysed respecting low or high yields with discriminance analysis /DA/.

MATERIAL AND METHOD

The Research Institute for Vegetable Oil and Detergent Industry collects each year 7o-loo variables from the sunflower fields of more than loo.ooo hectars.

Data of 17o fields - about 12.ooo hectars-were chosen to examine the effect of the variables on the yield. On each field the previous crop was the same - winter wheat - and the variety NSH Jugoslav hybrid was grown. So neither the big differences in soil tillage nor that in genetical basis had any significant influence on the result.

The environment and the plant technology were discribed by 18 quantitative variables selected from the originally collected 7o variables. The examination of more variables at the same time would make the analysis difficult to survey.

The applied method - the discriminance analysis - is suitable for separating two /or more/ groups on the basis of several quantitative variables in which the groups one by one are overlapping. The method is very similar to the multiple regression analysis but the dependent variable /Y/ is a qualitative instead of a quantitative one. The two variants are considered as two groups.

To analyse the field data of sunflower we applied the method in a new way. The 17o fields were divided into two groups and considered as the two groups of the DA.

group A: the yield is lower than the mean value /<2,1 t/ha/
group B: the yield is higher than the mean value />2,1 t/ha/
n_A = 81 fields n_B = 89 fields

So according to one variable /yield/ we divided the fields into two groups and then considering the 18 variables we determined the parameters of the discriminance function of the standardised differences and the correlation coefficients of the two groups.

With this function a complex value - called Z value - was calculated for every field without respecting if it is in the group of high or low yields. With this Z values the fields can be characterised instead of 18 variables only with one and it can be represented on a linear line.

The difference between the means of the two values of the groups gives the multivariate distance - the D^2 value:

$$D^2 = \bar{Z}_A - \bar{Z}_B$$

The importance of a given variable in respect to the classification into group A or B was determined with division of D^2. Considering D^2 as loo, the direct and indirect effects of the variables and the sum of them, as the total effect, can be established in %. The pozitive values separate the two groups, so they cause difference between the yields. The negative ones reduce the differences

between the groups because of the interactions.

RESULTS

Table 1. shows the 18 variables dealt with in DA, the mean values of them in the two groups and the differences and standardized differences between the mean values. It can be seen that the fields yielding more than the average yield, had 3,7 higher complex field value than the others yielding less. The potassium contents of the soil was 86,6 ppm higher in the soil of the high yielding fields.

Table 1. Basic table of discriminance analysis

Variables	Low yield A group $\overline{X}_A$	High yield B group $\overline{X}_B$	Difference $\overline{X}_A-\overline{X}_B$	Standard. difference
1 Previous crop t/ha	4,2	5,o	-o,816	-o,823
2 Compl. field val.	18,1	21,9	-3,729	-o,545
3 Soil texture K_A	45,7	4o,5	5,186	o,753
4 Humus %	2,7	2,4	o,249	o,3o6
5 Soil P_2O_5 ppm	174,5	175,2	-o,696	-o,oo8
6 Soil K_2O ppm	312,2	225,6	86,623	o,691
7 Sowing-time days	1o7,8	1o8,9	-1,o86	-o,188
8 Plant number looo	55,5	55,o	o,446	o,o83
9 Harvest-time days	267,3	265,1	2,18o	o,2o1
1o Oil %	46,5	47,3	-o,8o4	-o,373
11 looo grain weight	41,2	41,3	-o,o7o	-o,o12
12 Linol acid %	7o,o	7o,5	-o,5o5	-o,261
13 Total rainfall mm	256,4	279,7	-23,336	-o,5o2
14 N kg/ha fertilizer	84,o	89,9	-5,876	-o,195
15 P_2O_5 kg/ha fertilizer	97,9	99,1	-1,129	-o,o35
16 K_2O kg/ha fertilizer	114,o	141,8	-27,836	-o,595
17 Realized plant %	86,4	93,5	-7,o46	-o,65o
18 Realized crop %	78,1	111,2	-33,176	-1,959

The differences having dimension contain a lot of information for an agronomist but because of the different measurements it is difficult to compare them. That's why the last column of the table shows the differences in standardized form.

The variables having the most important differences between the means of the two groups can also be seen.

On Figure 1. the distribution of these variables in the high or low yielding groups are shown. According to the diagrams there is an important overlapping between the distribution of the two groups by variables. These are darkened on the figures.

Figure 2. shows the frequency distribution of Z values in two groups calculated by fields taking into consideration the 18 variables.

Figure 1. Distribution of the variables

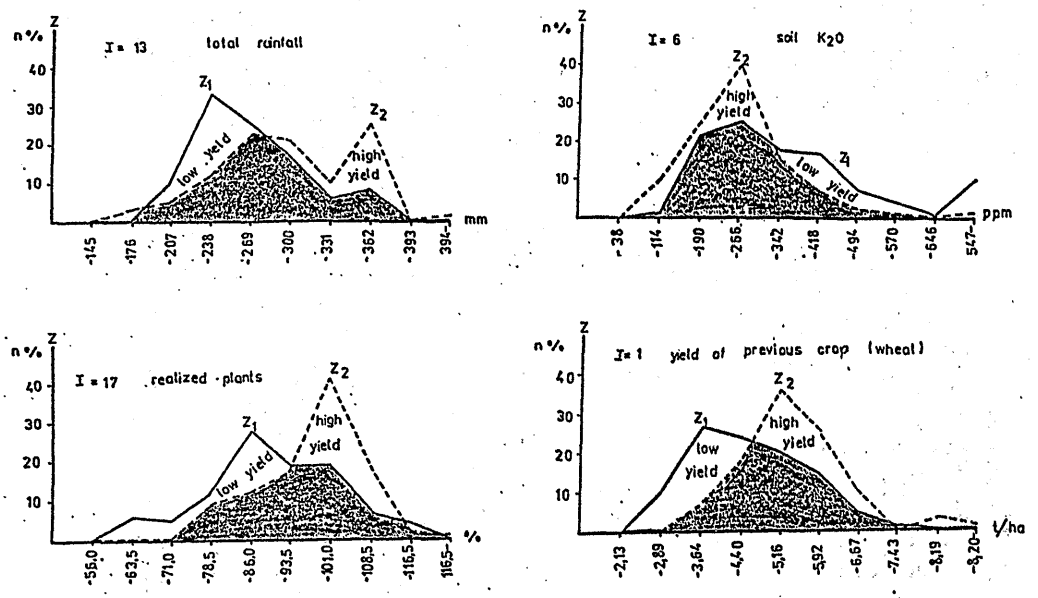
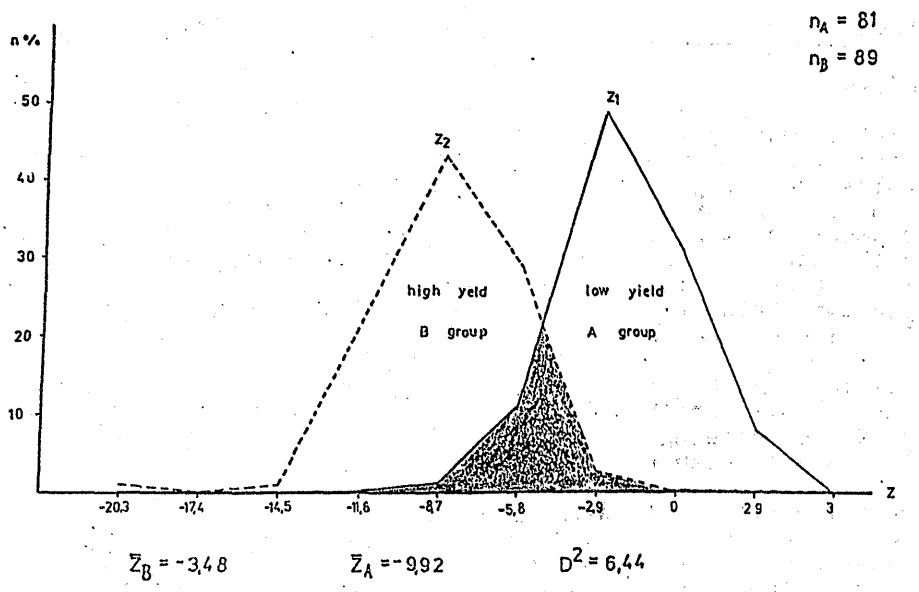


Figure 2. Distribution of the Z values of discriminance analysis



$\bar{Z}_B = -3{,}48$ $\bar{Z}_A = -9{,}92$ $D^2 = 6{,}44$

The overlapping on Figure 2. is much smaller than on the figures of Figure 1. so the two groups were separeted better taking 18 variables all together.

Table 2. shows the partition of D^2 according to the importance of examined variables. The table shows the total effect and its partition to direct and indirect effects. D^2 = loo %.

Table 2. Division of the D^2 % values calculated with discriminance analysis

Variables	Effects		
	total	direct	indirect
1 Previous crop t/ha	42,55	47,o5	-4,5o
2 Complex field value	7,45	3,42	4,o3
3 Soil texture K_A	19,25	11,33	7,92
4 Humus %	-1,o9	o,62	-1,71
5 Soil P_2O_5 ppm	o,oo	o,oo	o,oo
6 Soil K_2O ppm	12,42	6,21	6,21
7 Sowing-time days	-5,59	15,84	-21,43
8 Plant number looo	-2,79	18,17	2o,96
9 Harvest-time days	5,12	lo,24	-5,12
lo Oil %	2,79	1,o9	1,7o
11 looo grain weight	o,oo	6,21	-6,21
12 Linol acid %	-2,33	1,o9	-3,42
13 Total rainfall	7,45	4,o3	3,42
14 N kg/ha in fertilizer	-o,62	6,21	-6,83
15 P_2O_5 kg/ha in fertilizer	3,42	o,62	2,8o
16 K_2O kg/ha in fertilizer	-4,o4	4,o4	8,o8
17 Realized plants %	-6,21	26,71	32,92
18 Realized crop %	22,2o	68,32	46,12
Total:	loo,oo	loo,oo	

CONCLUSION

The discriminance analysis of the sunflower field data showed that the "yield of the previous crop" and the "realized yields %" variables had the greatest effect on the level of the yield. Both of them are complex values, they characterize the effects of the environment and the agrotechnics, as well. The high "realised yield %" actually shows the favourable coincidence of factors in sunflower growing. The "yield of the previous crop" - in the example the yield of the winter wheat - characterises the general level of the production.

Soil texture and K_2O ppm contents of the soil were also important. Very heavy soils whith high potassium contents are not recomended for sunflower growing. In 1985 - the year of the study - the more rainfall was favorable.

The negative effect of late sowing can be proved as an important indirect effect on the yield, as well. The increasing temperature may cause a very dry seedbed which results in a weak germination and a poor stand. These plants are more succeptible the diseases, etc. so finally the yield decreases. The significant importance of the required germination is proved by the effect of the "realised

plant number %" /number of harvested/planned plants/, as well. In this analysis the phosphate contents of the soil hadn't any effect on the yield because all the fields were well supplied with P.

Finally, it can be seen that the discriminance analysis is an adequate method to analyse the importance of a given factor in a multivariate system in agriculture.