MATHEMATICAL MODELS OF SOME INDICES IN SUNFLOWER

E.Pentchev,P.Ivanov.Institute for Wheat & Sunflower "Dobrudja"
G.Toshevo,Bulgaria

S U M M A R Y

Different statistical and númerical (quantative) methods are discussed in analyzing the relationships between some basis indices in sunflower.The features of the correlation and path - analysis are treated.Different numerical methods for approximation of the relationships existing concretely in the case of a variable quantity are applied.
The nature of the relationships between the studied indices is clarified.Their complex characteristic is proved.

INTRODUCTION

By the development of the computing engineering there is a possibility the methods of the matematical statistics and modeling to be automated and made popular because they are in the basis of the data analysis of different nature.We note that the insufficient and in complete knowledge of these methods can lead to their wrong application and produsing of false results.Partially this problem could be solved by the aid of specialized statistical program products conformed with the specifity of the objekt studied.In this paper we discuss the application of some mathematic - statistical methods for analysis of the relations between some essential indices in sunflower.

MATHERIAL AND METHODS

The experimental data are processed by the statistic program products "BIOSTAT",Version 2.1.,indicated for biological studies.The correlation coefficients on a phenotypic level are assessed by the aid of linear and curvilinear correlation analysis (1).The genetic correlation coefficients are calculated according to the model (2) :

1) $Rph=Rg + Re$,

where Rph - a correlation coefficients on a phenotypic level; Rg - correlation coefficient on a genetic level;Re assesses the influence of the ecological enviroment.
The correlation coefficients obtained in this way were subjected to a path-analysis (3) :

2) $\sum PiRij=Rik,\qquad i=1,k-1$

The investigation of the mathematic models was conducted by the application of a regression model of the type:

3) $F(x)=H(x) + Lh(x)$,

where $F(x)$ is real dependence;$H(x)$ is the approximate function nd $Lh(x)$ is the error in the approximation.The adequacy of the models is determined by the evaluation of the following statistical parameters:a mean quadratic deviation;standart error of the model and linear and curvilinear correlation coefficients (depending on the model) between the data obtained at the experiments and these simulated by the model. All computing procedures are realized in the "BIOSTAT",Version 2.1.It is written in Turbo Pascal,Version 6.0 in a dialogue regime.The program module for modeling offers a possibility for evaluation of the data by 13 dependiences occured most often in the nature and arrangements of the models with statistic criteria selected with possibilities for a graphic presentation and a prognosis for a given period.
Cultivars and hybrids tested in our country are studied as a part of network experiments organized by FAO.The experiments are carried out in two-year cycles.The hybrids are divided into two groups - early and medium late,42 being tested in the first group and 44 in the second one.

RESULTS AND DISCUSSION

The correlation analysis allows an investigation of the possible relationships between the indexes which are not bound to be causal,as the correlation coefficient is only a quantitative characteristics of the degree of dependence between two random quantities.Very often in a corelation coefficient tending to 0 a wrong conclusion is made for independence of the studied indices.The range of variation of the correlation coefficients on a phenotypic and genotypic level is given in Table1.The correlation coefficients between the indices seed yield and 1000 seed weight vary from - 0.317 to 0.564.In medium late hybrids they are positive,but are proved statistically only for 1984 (4).In early hybrids this relationship changes quantitatively and in direction.In 1982 and 1985 Rph and Rg are negative,although the rainfalls during these two years were quite different.Regarding this characteristics the first year is the best,while the second one is marked by the lowest values of the rainfalls (4).In yield and kernel size the correlation coefficients are one way positive only for the period 1984-1985.The picture is different also in the indices of yield and oil content of the kernel.The variation of Rph and Rg is more slight only in yield and protein content in the kernel confirming the tendency for a proven reverse dependence.In the rest indices the relationship is more reverse only between the indices oil and protein content in the kernel.The presence of an interaction genotype x ecological enviroment leads to the problem of representation of the excerption studied from random quantities of definite taxonomic group data.That is why in certain ecological conditions the correlation coefficient could indicate a proven relationship between two indices,and in others the lack of such relation.It results from this that correlation analysis provides a possibility for formation of specific hypothesis which are

T A B L E N.1

Variation of the correlation coefficients on a phenotypic and genotypic level

Indices	R(x,y)	2	3	4	5
1	Rph min.	-0.317	-0.224	-0.023	-0.726 c
	Rph max.	0.564 b	0.688 c	0.770 c	0.161
	Rg min.	-0.493 a	-0.300	-0.032	-0.294
	Rg max.	0.678 c	0.786 c	0.871 c	0.184
2	Rph min.		-0.386	-0.261	-0.484 a
	Rph max.		0.404	0.466 a	0.283
	Rg min.		-0.473 a	-0.676 c	-0.337
	Rg max.		0.635 c	0.694 c	0.286
3	Rph min.			-0.408 a	-0.685 c
	Rph max.			0.736 c	0.519 b
	Rg min.			-0.145	-0.726 c
	Rg max.			0.926 c	0.080
4	Rph min.				-0.956 c
	Rph max.				0.080
	Rg min.				-0.956 c
	Rg max.				-0.156

Key : 1-yield kg/dka
2-1000 seed weight.g
3-Kernel,%
4-oil in kernel %
5-protein in kernel,%

a --> P=0.95
b --> P=0.99
c --> P=0.999

valid only in certain ecological conditions (5).It is clear from the type of models (1) and (2)that the same conclusion is also valid for the genetic correlation coefficients and path coefficients.This affirmation is proved also by Table 2.The path - analysis is more reliable because it assesses the direct and indirect effect of the yield character.Table 2 shows path coefficients are different in size and direction in each of the indices studied.The conclusion is the study of these relationships only by the aid of the methods examined is risky and unreliable for the needs of the plant breeding.

T A B L E N.2

Variation of path - coefficients on a phenotypic and genotypic level

Indices	P(x,y)	2	3	4	5
1	Pph min.	-0.400	-0.600	-1.970	-2.080
	Pph max.	17.60	2.969	36.24	36.00
	Pg min.	0.033	-0.530	-0.330	-0.320
	Pg max.	0.480	0.340	1.300	0.684

Key :1-yield kg/dka
2-1000 seeds weight ,g
3-kernel,%
4-oil in kernel %
5-protein in kernel %

Specifying of the formed hypothesis can be done by the aid of the regression analysis (3).The analysis of the data by the "BIOSTAT" Version 2.1. pack proves on the basis of the statistical criteria the relationship between the yield and 1000 seed weight,kernel size and oil content in the kernel is approximated most adequately by the postulated model:

4) $E(y)=a + b\ lg(x)$,

where E(y) is the yield expected mathematically;a and b are real numbers.This confirms the hypothesis for a possible summarized model which describe the dependences studied,as in the cases of the different climatic conditions this model can assume an expression of another type.The graph of the logarithmic function gives the ground to prove that depending on the data obtained by experiments in the specific crop years the relation between the characters can attain also a linear type.The limiting effect of the 1000-seed weight,the kernel size and its oil content on the produktivity is proved also by this model,at least at the present sunflower plant stucture and seed chemical composition.

The results for the relation between the yield and protein content in the kernel are also specific.In this case the postulated models prove to be most suitable:

5) $E(y)=a + b/x$

and

6) $E(y)=1/(a+bx)$,

where E(y) is the yield expected,a and b are real numbers. The both models trace out a hyperbolic dependence from two different types.These models in different crop years can turn into linear ones depending on the location of the experimental data around the hyperbolic curve.The reverse propotional relationship between the yield and the protein content in the kernel are confirmed also by thise models.

The conducted analysis reveals the complex character of the relation between the yield with anyone of the indices by which the seeds are assessed,because of that the corelations and path coefficients vary highly depending on the enviroment factors.This proves once again the limited possibilities for an early evaluation of the breeding matherials and the necessity of a systematic and complex assesment of the pro - ductive possibilities of the hybrids tested.

CONCLUSIONS:

- The investigations on the yield relations with the morphological indices of the plants by the aid of a correla - tion and path analysis on a phenotypic and genotypic level show that wrong conclusions are possible by the results of these statistical criteria.
- The complex relationship between the yield and morpho - logical indices of the plant is proved by the application of the mathematic modeling.Equations for approximation of these dependences of a logarithmic and hyperbolic types are suggested.

REFERENCES:

1.Snedecor,G.,W.Cochran.Statistical methods.Jowa,1980.
2.Gentchev,G. et al.,Biometrical methods in the crop production,genetics and breeding.Sofia,1975.
3.Wright,S.Path-coefficients and Path-regression.Biometrics 16,1960.
4.Ivanov,P.Doctor's Dissertation.Sofia,1991.
5.Pentchev E.,I.Stoeva.Mathematical models of wheat quality indices.Sofia,Plant science 5,1989.